THE INTERACTIVE EFFECTS OF EXECUTIVE FUNCTIONING AND RELEVANT EXPERIENCE ON THEORY OF MIND DEVELOPMENT

by

Jeannette Elizabeth Benson

A thesis submitted to the Department of Psychology
In conformity with the requirements for
the degree of Doctor of Philosophy

Queen’s University
Kingston, Ontario, Canada
September, 2014

Copyright © Jeannette Elizabeth Benson, 2014
Abstract

This study investigated processes underlying the relation between executive functioning and false belief knowledge in preschool-aged children. The primary goal was to test the theory that executive functioning skills equip children with the cognitive tools necessary for experience-driven social conceptual change. To explore this possibility, 3.5-year-old children were recruited to participate in a longitudinal study that involved two study phases. During the initial phase, children were assessed on their false belief understanding and executive functioning skills, as well as abilities that served as relevant control variables (e.g., vocabulary skills). Two types of naturalistic experiences that relate to false belief development were also measured: parent mental state talk and having child-aged siblings. Six months later, families returned to the lab for the second phase of testing, and all measures were re-administered. Analyses examining concurrent relations among measures showed that at both study phases, there was an interactive effect of executive functioning skills and relevant experience on concurrent false belief task performance. This general pattern of results persisted—albeit in a weaker form—when controlling for consistency in executive functioning skills over time, age, vocabulary skills, and initial theory-of-mind knowledge. Longitudinal analyses revealed that there was an interactive effect of early executive functioning skills and experience on false belief understanding 6 months later, and these effects remained strong and significant when controlling for the relevant variables listed above. Moreover, the reverse longitudinal interactive effect—of later executive functioning and experience on early false belief knowledge—was not significant. Overall, effect sizes associated with the predicted longitudinal interaction were substantially larger than those associated with concurrent interactions. This study’s
longitudinal findings offer support for the hypothesis that executive functioning skills influence children’s ability to use relevant experience in the service of false belief concept development. Further, the concurrent pattern of findings suggests that executive functioning skills are also recruited for actively reasoning about other minds in the moment, when children are faced with false-belief-reasoning situations. Alternative theories are discussed, with emphasis placed on taking an integrated theoretical approach to effectively characterize the effects of executive functioning on false-belief reasoning.
Acknowledgements

First and foremost, I would like to extend my sincerest thanks to the families who generously volunteered their time to participate in this study. I would also like to thank the undergraduate students who assisted with this project, including Alexandra Cross, Dorothy Jenkins, Samantha Schoening, Autumn Tallick, and Hanna Wickstrom. In particular, I would like to thank Elizabeth McEwen for her contributions to this project as an honours student in our lab – her enthusiasm, hard work, and helpful insights were a valuable addition to this project.

I would also like to thank my fellow Early Experience Lab members for collectively creating such a friendly and supportive working environment. Special thanks to Sevda Bahtiyar, Martha Bailey, Nicole Bardikoff, Kaelen Boyd, Eric Bulmash, Anastasia Campbell, Mateya Dimnik, Samantha Drover, Monica Haberl, Annette Henderson, Chloe Hudson, Jennie Ito, Lindsay Kiernan, Cherie La Rocque, Chrissy Malyk, Teal Mackintosh, Haykaz Mangardich, Vanessa Schell, and Dustin Washburn. I feel very fortunate to have had the opportunity to work with such wonderful people over the years.

I am especially thankful to my advisor and mentor, Dr. Mark Sabbagh, for all of his guidance throughout my time as a graduate student. His support, feedback and encouragement have been invaluable, not only in the completion of this project, but also in the continued development of my writing style, research skills, and critical thinking abilities. I would also like to extend my thanks to the Developmental Psychology Department at Queen’s University for providing such a stimulating, supportive environment in which to learn. In particular, I would like to thank the members of my thesis committee, Dr. Elizabeth Kelley and Dr. Valerie Kuhlmeier, for their helpful feedback on this project, and for their kindness and enthusiasm throughout my time in the department.

I also wish to thank my family and friends for their continued love and encouragement. In particular, I would like to thank my aunt and uncle, Linda and Peter Hamstra, for their support as
fellow Kingstonians during my time at Queen’s. I also wish to extend my thanks to my boyfriend, Nick Pilon, for all of his love, encouragement, and comic relief over the years. His heart-felt support and balancing influence have made this achievement possible. Thanks also to the Pilon family in Mattawa, Ontario for welcoming me into their family with open arms.

Special thanks to my grandparents, Roy and Elsie Gorewich, and my stepfather, Dr. Stanley Hamstra, who have been central figures throughout my life and continue to provide me with unwavering love, support, and guidance. They have led by example, teaching me the importance of curiosity, ambition, and integrity. Words cannot express how thankful I am to have them as my role models. I am also thankful to my father, David Benson, for always encouraging me to strive for excellence, even when the going got tough. I know that he would be very proud of this accomplishment. Many thanks to Natalia Gontcharova for all of her kindness, encouragement, and support through very difficult times over the past few years. The completion of this project would not have been possible without her help.

Lastly, I would like to thank my mother, Dr. Nancy Benson, for all of her love, support, enthusiasm, and her ability to see the humour in almost any situation. Throughout my life she consistently encouraged me to reach for my dreams, and to strive for excellence. Above all, she taught me the importance of optimism, kindness, perseverance, and integrity. She has had a profound impact on almost every aspect of who I am as a person, and continues to be an inspiration in my life. This dissertation is dedicated to her loving memory.
# Table of Contents

Abstract .......................................................................................................................... ii
Acknowledgements ......................................................................................................... iv
List of Figures .................................................................................................................. ix
List of Tables ................................................................................................................... x
List of Abbreviations ...................................................................................................... xi

## CHAPTER 1: INTRODUCTION .................................................................................... 1

The Early Understanding Account .................................................................................. 4
Component Process Account ......................................................................................... 12
Emergence Account ....................................................................................................... 17
Summary ......................................................................................................................... 23
Goals of the Present Study ............................................................................................. 25
    Concurrent Analyses and Predictions ......................................................................... 27
    Longitudinal Analyses and Predictions ..................................................................... 29
    The Role of Age, Vocabulary Skills, and Initial Theory-of-Mind Knowledge ........... 30

## CHAPTER 2: METHOD ............................................................................................. 32

Participants ...................................................................................................................... 32
    Phase 1 ......................................................................................................................... 32
    Phase 2 ......................................................................................................................... 32
Materials and Measures ................................................................................................. 33
    False Belief Battery ...................................................................................................... 33
        Contents Change task ............................................................................................. 34
        Location Change task ............................................................................................ 34
        Appearance-Reality task ......................................................................................... 34
        Deceptive Pointing task ......................................................................................... 35
    Executive Functioning Tasks ...................................................................................... 36
        Bear/Dragon task ..................................................................................................... 36
        Grass/Snow task ..................................................................................................... 36
        Dimensional Change Card Sort (DCCS) task ......................................................... 37
<table>
<thead>
<tr>
<th>Measures of Relevant Experience</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parent mental state talk</td>
<td>38</td>
</tr>
<tr>
<td>Knowledge</td>
<td>41</td>
</tr>
<tr>
<td>Modulation of Assertion</td>
<td>42</td>
</tr>
<tr>
<td>Desire</td>
<td>42</td>
</tr>
<tr>
<td>Emotion</td>
<td>42</td>
</tr>
<tr>
<td>Pretend</td>
<td>43</td>
</tr>
<tr>
<td>Other</td>
<td>43</td>
</tr>
<tr>
<td>Number of Child-Aged Siblings in the Home</td>
<td>44</td>
</tr>
<tr>
<td>Vocabulary Measure</td>
<td>44</td>
</tr>
<tr>
<td>Theory of Mind Scale</td>
<td>44</td>
</tr>
<tr>
<td>Diverse Desires Task</td>
<td>45</td>
</tr>
<tr>
<td>Diverse Beliefs Task</td>
<td>45</td>
</tr>
<tr>
<td>Knowledge Access Task</td>
<td>45</td>
</tr>
<tr>
<td>Procedure</td>
<td>46</td>
</tr>
<tr>
<td>Phase 1</td>
<td>46</td>
</tr>
<tr>
<td>Phase 2</td>
<td>46</td>
</tr>
</tbody>
</table>

### CHAPTER 3: RESULTS

Performance on Measures Independently | 49
False Belief Battery                | 49
RC-EF Battery                       | 52
Bear/Dragon Task                    | 52
Grass/Snow Task                     | 52
DCCS Task                           | 52
Relations Among Tasks and Scale Aggregating | 53

Measures of Relevant Experience     | 57
Parent Mental State talk            | 57
Child-Aged Siblings                 | 58
Experience Aggregate                | 58
List of Figures

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Early understanding, component process, and emergence accounts of the relation between RC-EF and false belief reasoning</td>
<td>5</td>
</tr>
<tr>
<td>2. Histogram and Q-Q Plot Illustrating RC-EF Aggregate Scores in the Full Sample at Phase 1</td>
<td>54</td>
</tr>
<tr>
<td>3. Histogram and Q-Q Plot Illustrating RC-EF Aggregate Scores in the Returning Subset at Phase 1</td>
<td>55</td>
</tr>
<tr>
<td>4. Histogram and Q-Q Plot Illustrating RC-EF Aggregate Scores in the Returning Subset at Phase 2</td>
<td>56</td>
</tr>
<tr>
<td>5. Histogram and Q-Q Plot Illustrating Experience Aggregate Scores in the Full Sample at Phase 1</td>
<td>60</td>
</tr>
<tr>
<td>6. Histogram and Q-Q Plot Illustrating Experience Aggregate Scores in the Returning Subset</td>
<td>61</td>
</tr>
<tr>
<td>7. Two-Way ANOVA Showing the Concurrent Interaction between RC-EF and Relevant Experience in Predicting False Belief Understanding at Phase 1</td>
<td>67</td>
</tr>
<tr>
<td>8. Two-Way ANOVA Showing the Concurrent Interaction between RC-EF and Relevant Experience in Predicting False Belief Understanding at Phase 2</td>
<td>68</td>
</tr>
<tr>
<td>9. Two-Way ANOVA Showing the Longitudinal Interaction between Phase 1 RC-EF and Relevant Experience, Predicting Phase 2 False Belief Understanding</td>
<td>71</td>
</tr>
<tr>
<td>10. An Integrated Model of the Mechanisms Underlying the Relation between RC-EF skills and False Belief Reasoning in Preschool-Aged Children</td>
<td>84</td>
</tr>
</tbody>
</table>
## List of Tables

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. A Summary of the Predictions made by the Component Process and Emergence Accounts</td>
<td>29</td>
</tr>
<tr>
<td>2. Categories and Examples of Mental State Terms</td>
<td>41</td>
</tr>
<tr>
<td>3. Task Administration</td>
<td>48</td>
</tr>
<tr>
<td>4. Descriptive Statistics of Measures within Full Sample at Phase 1</td>
<td>50</td>
</tr>
<tr>
<td>5. Descriptive Statistics of Measures within Returning Subset at Phases 1 and 2</td>
<td>51</td>
</tr>
<tr>
<td>6. Inter-Correlations Among Key Measures in Full Sample at Phase 1</td>
<td>64</td>
</tr>
<tr>
<td>7. Inter-Correlations Among Key Measures in Returning Subset at Phases 1 and 2</td>
<td>65</td>
</tr>
<tr>
<td>8. Results of Two-Way ANCOVAs Controlling for Age, Theory-of-Mind Scale Scores, and PPVT Scores</td>
<td>74</td>
</tr>
</tbody>
</table>
# List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADHD</td>
<td>Attention Deficit/Hyperactivity Disorder</td>
</tr>
<tr>
<td>FB</td>
<td>False Belief</td>
</tr>
<tr>
<td>PPVT-IV</td>
<td>Peabody Picture Vocabulary Test, Fourth Edition</td>
</tr>
<tr>
<td>RC-EF</td>
<td>Response-Conflict Executive Functioning</td>
</tr>
<tr>
<td>ToM</td>
<td>Theory of Mind</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

“Understanding changes in children’s thinking presents all of the demands of understanding their thinking at any one time, plus the added demands of understanding what is changing and how the change is being accomplished.” (Siegler, 1995, p. 225).

‘Theory of Mind’ refers to the understanding that others have internal mental states that both influence and are influenced by human behaviour (Wellman, 1990). Over the past 30 years, the development of theory-of-mind knowledge has received empirical attention from a diverse range of researchers, including social, cognitive, comparative, developmental, and clinical psychologists. The resulting body of work shows that transitions in understanding a range of socially relevant constructs—including beliefs, desires, intentions and emotions—occur over the preschool years (e.g., Wellman & Liu, 2004). Cross-cultural research on theory-of-mind suggests some level of variability in the developmental timing of theory-of-mind acquisition across cultures (e.g., Wellman, Fang, Liu, Zhu, & Liu, 2006). Nevertheless, among typically-developing children, research suggests surprising regularity, both in the order and age at which children come to understand mental concepts (Callaghan et al., 2005; Peterson, Wellman, & Liu, 2005; Wellman, 2002; Wellman & Liu, 2004).
Within the research on theory-of-mind development, significant emphasis has been placed on studying children’s understanding that beliefs can be false. One standard task designed to measure false belief knowledge involves recounting a story to children using dolls and toy props. Children are shown a character hide an object in one location and leave the scene. In her absence, a second character hides the object in an alternative location, and children are asked where the first character will search for her object upon returning. To answer correctly, children must understand that the first character now has a false, outdated belief about her object’s whereabouts, and will search for it accordingly in the original (now empty) location.

When children reveal false belief knowledge by succeeding on these kinds of tasks, they are demonstrating the understanding that mental states are subjective representations of the world, and as such, have the potential to be inaccurate (Wellman, Cross, & Watson, 2001). The mental state concepts that typically come to be understood prior to false belief (e.g., desire) do not require children to navigate and understand a direct conflict between mental representations and reality (Perner, 1991). Thus, false belief task performance is often considered to be the litmus test for a representational theory of mind. A wealth of research indicates that children do not tend to demonstrate consistently correct performance on standard false belief tasks until 4 to 5 years of age (e.g., Wellman et al., 2001). Further research shows that the preschool shift in performance on standard false belief tasks is correlated with just the kinds of naturally-occurring social behaviours that one would expect to be dependent on mature false belief reasoning, including pretending (Taylor &

A significant body of research has been directed at understanding the cognitive and experiential factors associated with the preschool shift in false belief understanding. One such factor is the collection of domain-general abilities associated with frontal lobe activity known as executive functioning skills. Executive functioning skills encompass the processes that underlie goal-directed behaviour, including self-regulation, planning, working memory, response inhibition, and resistance to interference (Zelazo, Carlson, & Kesek, 2008; Zelazo & Müller, 2002). Children’s performance on false belief tasks correlates most consistently with performance on response-conflict executive functioning (RC-EF) tasks, which involve inhibiting a dominant or prepotent response option in order to respond correctly (e.g., in the game ‘Simon Says’) (e.g., Carlson & Moses, 2001). The correlation between false belief performance and RC-EF skills typically remains sizeable, even after more general factors—including age, sex, and language skills—are statistically controlled.

While the relation between RC-EF skills and false belief task performance is well established, the mechanisms underlying this relation remain the subject of debate (Benson & Sabbagh, 2009). In the following paragraphs, I outline three general theories on this topic, and present relevant research findings for each.
The Early Understanding Account

Some researchers have proposed that even infants possess underlying false belief knowledge, but their immature RC-EF skills prevent them from demonstrating that knowledge in many experimental settings (e.g., Leslie, 2005). From this account, the relation between RC-EF and preschoolers’ false belief task performance exists because RC-EF skills are necessary for children to express their underlying false belief knowledge (see Figure 1a). One reason this may be true is that the surface structure of false belief tasks poses significant executive demands over and above the demands placed on false belief reasoning, and children require sufficiently developed RC-EF skills to overcome those demands. For instance, when children are prompted to make an active response in false belief task scenarios, they are challenged to refer to a location where something is not. This unusual response is thought to be challenging given our habitual tendency to indicate where objects actually are; RC-EF skills may be necessary to suppress or override this tendency in order to respond according to false beliefs. From the early understanding account, the preschool shift in false belief task performance is driven by the maturation of executive functioning skills, and not by underlying shifts in social conceptual knowledge.
Figure 1. Early understanding, component process, and emergence accounts of the relation between RC-EF and false belief reasoning.
Evidence for this account comes from studies showing that while children do not pass standard false belief tasks until late preschool, infants and young children show knowledge of false-belief-based behaviour in looking-time and preferential looking studies that do not require an active response (e.g., Onishi & Baillargeon, 2005). For instance, Onishi and Baillargeon (2005) showed 15-month-old infants non-verbal false belief scenarios that were similar in structure to the standard false belief task described above. Following the standard location-change events, infants were shown one of two possible outcomes: either the agent searching for her object in its previous, now-empty location (i.e., false-belief-consistent search behaviour), or the agent searching in the object’s current location (i.e., false-belief-inconsistent search behaviour). Results from these trials along with a series of relevant control conditions indicate that infants look longer when actors ultimately search in locations that are inconsistent with their beliefs. This suggests that infants have expectations—at least on some level—that agents will act in false-belief-consistent ways.

Further evidence for this account comes from the finding that when preschoolers are presented with standard false belief tasks like the one described above, they do not respond to the forced-choice test question at chance levels, instead consistently referring to the object’s current location. Although there are other ways to interpret this finding, according to this account, children must inhibit the prepotent urge to reference reality (e.g., point to the object’s current location), and it is this ability that is under-developed in young preschool-aged children.
Another piece of evidence for this account is that children tend to demonstrate more advanced false belief knowledge when the RC-EF demands of standard tasks are lowered (Carlson, Moses, & Hix, 1998; Couillard & Woodward, 1999; Freeman, Lewis, & Doherty, 1991; Mitchell & Lacohée, 1991; Moses, 1993; Wellman & Bartsch, 1988; Zaitchik, 1990). For example, if the standard task is modified so that the hidden object is absent at the time when the test question is posed (e.g., while the character is gone, her chocolate is removed and then eaten), then young preschoolers’ performance improves. Performance also improves when children are asked to respond in a way that is novel and thus does not have a prepotent response (i.e., with a sign, Carlson et al., 1998; Couillard & Woodward, 1999). These findings suggest that when the demand to inhibit one’s usual tendency to point to an object’s true location is decreased, children are more successful at demonstrating false belief understanding on standard tasks.

Further evidence for this account comes from Friedman and Leslie (2005), who showed that increasing the executive demands of false belief tasks results in a decline in performance among preschool-aged children. In one study, children were shown the typical events involved in standard false belief tasks – a character hid an object, left the scene, and the object was moved to an alternate location in his absence. However, children were told that the character wishes to avoid the object, rather than find it. Thus, children needed to inhibit the more typical assumption that the character will move towards the object, rather than away from it. Results indicated that 4-year-olds who were successful on standard false belief tasks were unable to pass this modified task, suggesting that the increased executive
demands placed a limit on their ability to demonstrate underlying false belief knowledge (Cassidy, 1998; Leslie, German, & Polizzi, 2005).

Finally, neuroimaging work shows that when false belief tasks are given to adults, regions associated with both mental-state reasoning and executive functioning processes are activated (Saxe, Schulz, & Jiang, 2006). According to proponents of the early understanding account, these findings suggest that there are indeed executive processing demands inherent to false belief tasks. It follows that young children’s immature RC-EF skills may prevent them from revealing underlying false belief knowledge on these measures.

Taken together, this research highlights two points: 1) that some false-belief-relevant knowledge is available, even in infancy, and 2) that standard false belief tasks pose significant surface demands on RC-EF skills. Nevertheless, further research suggests two important counter-points. The first is that under-developed RC-EF skills do not appear to be masking otherwise developed, mature false belief reasoning abilities in young preschool-aged children. Evidence for this comes from a meta-analysis by Wellman and colleagues (2001) showing that while lowering the RC-EF demands inherent to false belief tasks does raise young preschoolers’ performance significantly, it does not bring performance to above-chance levels. Thus, young preschoolers do not demonstrate a systematic, mature understanding of false beliefs, even when the superficial RC-EF demands of false belief tasks are lowered. This finding calls into question the early understanding account’s hypothesis that young preschool-aged children have underlying, mature false belief
knowledge that is masked by their inability to negotiate surface task demands.

A second counter-point to the early understanding account is that the relation between RC-EF skills and false belief task performance appears to be more complex than originally hypothesized, existing independent of the surface task demands described above. One line of evidence for this claim comes from studies showing that preschoolers’ performance on tasks that reduce the need to inhibit a prepotent response continue to correlate with RC-EF skills (e.g., Perner & Lang, 1999; Perner, Lang, & Kloo, 2002). For instance, Perner and colleagues employed a version of the standard false belief task in which children were ultimately shown the outcome of the character’s false belief (i.e., they were shown the character searching for his object in the out-of-date location). Children were then asked to explain the character’s actions after the fact. Unlike standard false belief tasks, this task did not require children to inhibit the urge to point to an object’s true location; instead, they were asked to explain the character’s search behaviour after having seen where he would look for his object. Nevertheless, performance was strongly correlated with RC-EF skills, and this relation remained strong and significant when age, verbal intelligence, and performance on control questions were statistically controlled.

A more complex relation between RC-EF and false belief is also suggested by findings from cross-cultural research. This work has identified cultures in which preschool-aged children show an advanced timetable of RC-EF development in the absence of similar advances in false belief task performance (Oh & Lewis, 2008; Sabbagh, Xu, Carlson, Moses, & Lee, 2006). For instance, in a study by Sabbagh and colleagues (2006), Chinese
3.5-year-olds were performing at the level of U.S. 4.0-year-olds on RC-EF tasks, yet the two groups performed similarly on false belief tasks. Nevertheless, within the group of Chinese preschoolers, RC-EF remained predictive of individual differences in performance on false belief tasks. This is compelling evidence that attaining a particular level of RC-EF skills is not alone sufficient to improve false belief task performance. Thus, RC-EF skills appear to be necessary but not sufficient for false belief reasoning.

Atance et al. (2010) provide further evidence against straightforward predictions of the early understanding account by examining preschool-aged children’s response latencies on false belief tasks. According to the early understanding account, children automatically engage in belief reasoning in false belief task scenarios by way of innate or early-developing social-cognitive processes. However, to respond correctly, they must subsequently engage in the process of overcoming RC-EF-based challenges associated with false belief reasoning scenarios. From this account, it is this additional step that develops over the preschool years and leads to improved performance on false belief tasks. It follows that correct responses on false belief tasks should have longer response latencies than incorrect responses, as answering correctly should require time to engage in an extra RC-EF-based processing step. Atance and colleagues showed that while 3.5-year-old children’s correct false belief task responses occurred more slowly than their incorrect ones, this pattern was reversed in older children, with 4.5- and 5.5-year-olds’ correct answers tending to occur more quickly than their incorrect answers. These findings suggest that the early understanding account does not fully capture developmental shifts in false belief task
performance over the preschool period.

Finally, longitudinal research findings support a unidirectional relation between RC-EF skills and false belief development over time (Carlson, Mandell, & Williams, 2004; Flynn, 2007; Hughes, 1998). Studies show that RC-EF skills predict later false belief understanding across periods ranging from 5 months to 1 year, even when age, verbal ability, and initial false belief performance are controlled. In contrast, early false belief understanding does not similarly predict later RC-EF skills. One potential confound in these studies relates to consistency in RC-EF skills over time; to the extent that early RC-EF skills are a good predictor of later RC-EF skills, then any longitudinal relations between early RC-EF skills and later false belief understanding may actually be driven primarily by concurrent relations between variables. Although few studies have taken this possibility into account in their analyses, Carlson et al. (2004) did so by showing that initial RC-EF skills were predictive of false belief skills 10 months later, even when controlling for later RC-EF skills. These findings provide evidence for a unidirectional, longitudinal relation between early RC-EF skills and later false belief understanding. According to the early understanding account, there is no obvious reason to expect this finding; if RC-EF skills simply allow children to overcome superficial task demands in order to reveal their underlying knowledge when faced with false belief tasks, then concurrent but not longitudinal relations among measures would be expected. These findings provide yet another piece of evidence to suggest that the relation between RC-EF skills and false belief may be more complicated than hypothesized by proponents of the early understanding
In sum, although research shows that titrating RC-EF demands affects false belief task performance in predictable ways, it does not appear that immature RC-EF skills are preventing young preschool children from demonstrating otherwise mature false belief understanding in experimental settings. Further, it does not appear that the surface executive demands inherent to false belief tasks can account for the full developmental association between variables. Instead, research suggests a complex pattern of associations between RC-EF and false belief task performance over the preschool years that do not align with the early understanding framework.

**Component Process Account**

A second possible account for the relation between RC-EF and false belief task performance is one that I have chosen to label the ‘component process account’. From this account, RC-EF skills play an intrinsic role in the process of actively reasoning about other minds in false belief contexts. According to Samson and colleagues (Samson, Apperly, Kathirgamanathan, & Humphreys, 2005), false belief reasoning in most contexts can be broken down into two component processes: 1) overcoming the tendency to focus on one’s own perspective (deemed ‘self-perspective inhibition’), and 2) inferring the actual content of another individual’s belief (deemed ‘other perspective taking’). From this account, mature RC-EF skills exert their effect on false belief reasoning by enabling self-perspective inhibition in the service of focusing attention on other perspectives (see Figure 1b).
The idea that one’s self-perspective has particular cognitive salience is a common one in the literature. Similar theories have been put forth by a range of social-cognitive researchers, who have referred to the role of ‘a reality bias’ (Mitchell & Lacohée, 1991), ‘epistemic egocentrism’ (e.g., Royzman, Cassidy, & Baron, 2003), ‘hindsight bias’ (e.g., Bernstein, Atance, Meltzoff, & Loftus, 2007), or ‘the curse of knowledge’ (e.g., Birch & Bloom, 2007) in reasoning about other minds. All of these theories center on the notion that shifting the focus from one’s current mental perspective onto the differing perspective of someone else is a challenging process that is necessary for successful mental state reasoning.

Importantly, self-perspective inhibition and other perspective taking need not always co-occur; it is possible (although presumably uncommon) to encounter circumstances in which individuals realize another person’s belief is false when they themselves have no specific belief on a given topic, thus limiting the need for self-perspective inhibition. However, most instances of false belief reasoning – including those involved in standard false belief tasks—do require self-perspective inhibition; false beliefs are typically recognized in circumstances in which we are privy to facts that signal an opposing belief to be inaccurate.

Intriguing evidence for this dissociation between component processes involved in false belief reasoning comes from work by Samson et al. (2005), who examined the false belief task performance of a patient with right prefrontal and temporal brain damage following a stroke (referred to as patient “WBA”). This damage led to impairments in
WBA’s executive functioning abilities—in particular, his RC-EF skills. Results showed that patient WBA was impaired on standard false belief tasks, with his answers in line with those of a young 3 year old. In contrast, WBA showed remarkably strong performance on modified false belief tasks that circumvented the need for self-perspective inhibition. In these tasks, WBA was told to locate an object with a woman’s assistance. He was then shown a video in which a woman looked on while a man placed an object into one of two identical boxes, the outcome of which remained hidden from WBA. The woman temporarily left the room, and in her absence the boxes were swapped. Upon returning, the woman (who now held a false belief) pointed to one of the boxes to indicate the object’s whereabouts. Because WBA did not see which box the object was placed in at the beginning of the scenario, he did not have a competing belief about the object’s whereabouts to ignore when it came time to reason about the woman’s belief. WBA’s success on these tasks in conjunction with his impairment on standard false belief tasks provide evidence that RC-EF skills may indeed be recruited for self-perspective inhibition during false belief reasoning scenarios.

The component process account bears similarity to the early understanding account in stipulating that RC-EF skills are recruited for performance on standard false belief tasks. It also suggests that a lack of RC-EF skills may obscure measurement of other aspects of false belief reasoning (i.e., other perspective taking); according to the component process account, if children do not have sufficient RC-EF skills, then in most false-belief-reasoning
contexts they will be unable to avoid focusing on their own perspective in order to reason about others’ beliefs.

Despite these commonalities, the component process account differs from the early understanding account in three important ways. First, it suggests that self-perspective inhibition is a component process involved in false belief reasoning, rather than a separate process that has the potential to mask true false belief reasoning. Second, the theories differ in the point at which RC-EF skills are thought to be involved; from the component process account, RC-EF skills play an integral role in making other perspective taking possible, at least in most contexts. In contrast, proponents of the early understanding account typically argue that self-perspective inhibition is necessary after other perspective taking has occurred, at the point when children are prompted to respond on standard false belief tasks (Baillargeon, Scott, & He, 2010; Leslie & Polizzi, 1998). Finally, unlike the early understanding account, the component process account does not suggest that false-belief relevant concepts are innate or early emerging; instead, it leaves open the possibility that these concepts develop over the preschool years. Nevertheless, the component process account does not specifically address the factors involved in social conceptual change, focusing instead on the role played by RC-EF skills in online false belief reasoning.

In comparison with the early understanding account, the component process account appears to be compatible with a wider range of the accrued findings on the relation between false belief understanding and RC-EF skills. For instance, consider research showing that performance on explanation false belief tasks continues to correlate
with RC-EF skills. Although explanation tasks alleviate the need to point to an empty location, they still require children to inhibit their own perspective to reason about another individual’s false belief. That is, to generate a reasonable answer for why a character is searching for his object in an empty location, children must inhibit the tendency to focus on their own, accurate knowledge about the object’s whereabouts. The cross-cultural findings described above are also compatible with this general theory. Although RC-EF skills are hypothesized to be necessary for self-perspective inhibition, they are not alone sufficient for success on false belief tasks. Thus, despite 3.5-year-old Chinese children’s advanced RC-EF skills, other factors may have limited their performance on false belief tasks.

Nevertheless, the component process account does not straightforwardly explain findings in support of a longitudinal relation between early RC-EF skills and later false belief task performance (e.g., Carlson et al., 2004). If RC-EF skills are recruited for self-perspective inhibition during false belief reasoning, then we would expect false belief task performance to specifically relate to concurrent levels of RC-EF skills. There is no obvious reason to expect relations between false belief performance and one’s RC-EF skills several months prior to false belief task administration, at least not independent of concurrent RC-EF effects. That is, if RC-EF skills begin to exert their effect on false belief reasoning at the moment when a false belief reasoning context presents itself (at which point the need for self-perspective inhibition arises), then it should be RC-EF skills
in the moment and not RC-EF skills at another point in time that predict false belief task performance. In the following section I review a third account of the relation between RC-EF skills and false belief understanding that offers a possible explanation for these longitudinal findings.

Emergence Account

The early understanding and component process accounts suggest that RC-EF skills are necessary for successful performance on false belief tasks, either by enabling the negotiation of superficial task demands, or by facilitating sub-component processes involved in online false belief reasoning. The ‘emergence’ account takes a different perspective, suggesting that RC-EF skills contribute to the underlying development of false belief concepts in young children. According to this account, RC-EF skills facilitate the process of developing conceptual understandings of other minds from relevant experience (see Figure 1c).

This account is consistent with the hypothesized role of RC-EF skills in conceptual development within a number of domains, including mathematics and language (e.g., Blair & Razza, 2007; Bull & Scerif, 2001; De Beni, Palladino, Pazzaglia, & Cornoldi, 1998; Espy et al., 2004; McClelland et al., 2007). For instance, Espy and colleagues (2004) showed that inhibitory control skills predicted mathematical abilities in preschool-aged children, over and above the effects of age, vocabulary, and mathematical education. The prevailing explanation for these findings is that RC-EF skills play a role in developing
mathematical and linguistic knowledge over time. The emergence account suggests that RC-EF skills are important for social conceptual change, as well.

The emergence account aligns well with the ‘theory’ theory account of theory-of-mind development, which is a framework put forth to characterize and explain changes in mental state knowledge over the preschool period (Gopnik & Wellman, 1994). According to this framework, children formulate a theory to explain the actions of agents in their surroundings. They begin with non-representational understandings of a restricted range of mental constructs (e.g., desires) that are used to both explain and predict human behaviour. As children encounter evidence that runs counter to their expectations over time, they make revisions to their underlying explanatory framework and ultimately develop a mature theory of mind that includes subjective, representational beliefs. The emergence account adds to the theory theory by suggesting that RC-EF skills play some role in enabling this process of underlying theory change.

Unlike the early understanding account, both the emergence and component process accounts are compatible with the notion that false-belief-relevant concepts develop over the preschool years, and make subsequent false belief reasoning possible. Both are also compatible with the notion that conceptual developments are driven by exposure to relevant experiences from which to learn, although the component process account makes no definitive predictions in this regard. Where these two theories differ is in the role that RC-EF skills play in false belief reasoning; while the emergence account stipulates an interaction between RC-EF and relevant experience in the process of social conceptual
change over time, the component process account makes no similar conjecture, suggesting that RC-EF skills instead exert their effect by facilitating self-perspective inhibition during online false belief reasoning.

An important question that follows from the emergence account concerns the specific means through which RC-EF skills exert their effect on experience-driven social conceptual development. RC-EF skills may contribute to false belief concept development in at least two general ways. First, children with advanced RC-EF skills may be better able to effectively engage in social interactions, which themselves provide relevant experiences from which to develop a theory of mind (Flynn, 2007; Hughes, 1998). Second, once engaged in social interactions, RC-EF skills may be necessary for children to develop conceptual understandings from relevant feedback. RC-EF skills may have this facilitative effect on learning from experience by enabling children to 1) consider and attend to relevant information, 2) reflect upon discrepancies between theory-driven expectations and subsequent outcomes, and 3) flexibly update previously-established theories.

According to the emergence account, RC-EF skills work in conjunction with exposure to relevant experiences in order to facilitate false belief development. Thus, critical to this account is the notion that false belief development rests not only on RC-EF skills, but also on exposure to information about other minds. Indeed, research has shown that children’s false belief task performance can be improved through training methodologies designed to provide a concentrated dose of false-belief-relevant experience (e.g., Slaughter & Gopnik, 1996). Further research suggests that theory-of-mind
development is related to a number of socially-relevant environmental factors. For instance, research shows that having siblings is positively related to preschoolers’ false belief reasoning (e.g., Perner, Ruffman, & Leekam, 1994; Ruffman, Perner, Naito, Parkin, & Clements, 1998). This effect exists specifically for child-aged siblings and not for infants and toddlers, who lack the language skills necessary for the kinds of rich social interactions that are thought to be important for learning about other minds. Further research shows that children’s false belief understanding is uniquely related to parents’ use of words that explicitly highlight the subjective nature of beliefs, including knowledge terms (e.g., think, know), and terms that modulate the certainty of beliefs (e.g., might, probably, certainly; Ruffman, Slade, & Crowe, 2002). In contrast, false belief understanding is not similarly related to talk that focuses on other mental processes, such as emotions and desires. The prevailing theory that has emerged from this work is that experiences that are relevant to false beliefs play an important, facilitative role in the development of children’s false belief knowledge.

Similar to the component process account, the emergence account offers a compatible framework for understanding many of the study results described above. First, it is compatible with finding concurrent relations between RC-EF skills and false belief task performance. Research suggests relative consistency in RC-EF skills over the preschool period, with children’s early RC-EF skills predicting individual differences in performance years later (e.g., Hughes & Ensor, 2007). It follows, then, that concurrent relations between measures would likely exist as a consequence of any prolonged, longitudinal effects of RC-
EF skills on false belief conceptual development.

The emergence account is also compatible with the finding that RC-EF skills correlate with performance on false belief tasks designed to have lowered executive demands; if RC-EF skills are important for the development of underlying false belief knowledge, then we would expect RC-EF skills to be correlated with performance on any tasks that tap an understanding of false beliefs, regardless of whether the surface characteristics of the task require RC-EF skills.

This account also provides an explanation for cross-cultural studies showing that advanced RC-EF skills do not always translate to advances in false belief performance. According to this theory, RC-EF skills work in conjunction with exposure to relevant experience in the process of false belief development. Thus, while children from some cultures may show advanced RC-EF skills relative to others, it is possible that they have comparably less exposure to the kinds of experiences upon which false belief development depends (Sabbagh et al., 2006). Indeed, research points to significant cross-cultural variability in the types of experiential factors that typically correlate with false belief development. For instance, number of child-aged siblings living in a child’s household is significantly related to preschoolers’ false belief understanding (e.g., Perner et al., 1994; Ruffman et al., 1998), presumably because they offer exposure to other mental perspectives. As a result of government-enforced laws prohibiting more than 1 child per household in China, Chinese children may not have as many opportunities to engage with siblings when compared to children in other parts of the world, thus reducing their levels of
false-belief-relevant experiences.

In contrast to the early understanding and component process accounts, the emergence account offers a clear explanation for research showing a longitudinal relation between early RC-EF skills and later false belief understanding. The emergence account suggests that progressive RC-EF-dependent theory change takes place over the span of months, ultimately forming the understandings that are important for later false belief performance. From this account, then, we would expect RC-EF skills in the early preschool period to strongly predict subsequent advances in false belief task performance.

In a recent study, Benson, Sabbagh, Carlson, and Zelazo (2013) tested the interactive effects of RC-EF and experience on false belief development by examining whether individual differences in preschoolers’ RC-EF skills predicted benefits from a false belief training regimen. Given the obvious challenges associated with experimentally eliciting fast, widespread changes in a construct that typically takes months to develop, the training regimen was designed to advance one specific aspect of false belief knowledge, namely that people will search unsuccessfully for an object in cases where the object has been unexpectedly moved. Following an initial assessment of false belief understanding, children participated in 2 training sessions over a 2-week period during which standard false belief task scenarios were presented and explained. Children’s false belief understanding was then re-assessed, providing a direct measure of false belief improvement over the testing period. Results of this study showed that RC-EF skills strongly and consistently predicted improvements in false belief task performance. Children with higher
RC-EF skills not only showed more improvements in false belief task performance over the testing period, they also realized those improvements sooner, and provided explanations for their answers that suggested they were more successful at using the training feedback to advance their knowledge. These relations remained strong, even when age, language abilities, initial theory-of-mind knowledge, and concurrent improvements in RC-EF skills were included as control variables in the analyses.

Findings from this study were presented as evidence for the emergence account of false belief development. However, they may also be compatible with the component process theory; because this training study took place over a short 2-week period, it was not possible to tease apart the effects of RC-EF skill levels during concept acquisition from those ultimately available to children during subsequent false belief task performance. It is therefore possible that training experience (and not RC-EF skills) caused advances in children’s underlying false belief concepts, with only the high RC-EF children able to demonstrate those advances on standard false belief tasks. Thus, research has yet to gain direct evidence for the central premise of the emergence account; i.e., that RC-EF skills interact with experience to contribute to underlying false belief concept formation over time.

Summary

Research has identified a relation between RC-EF skills and false belief development, and I have outlined three different proposed accounts of the mechanisms
underlying this relation. Each of these accounts offers a different hypothesis regarding how best to characterize the roles of both RC-EF skills and conceptual development in false belief reasoning.

According to the early understanding account, RC-EF skills are necessary for children to reveal underlying mental state knowledge on standard false belief tasks. This account hypothesizes that false belief concepts are innate or early emerging, but that executive demands inherent to standard false belief tasks prevent young children from demonstrating such knowledge in research contexts. From this account, advances in false belief task performance over the preschool years are driven primarily by the maturation of executive functioning skills. In line with this account, research suggests that standard false belief tasks do indeed involve RC-EF demands. However, it does not appear that these demands serve to mask otherwise mature false belief knowledge. Moreover, these executive demands do not appear to fully account for the observed relation between RC-EF abilities and false belief reasoning.

According to the component process account, RC-EF skills are important for the process of self-perspective inhibition during false belief reasoning. Self-perspective inhibition is characterized as a necessary precursor to effectively generating the content of other minds in most false belief reasoning contexts. This account leaves open the possibility that underlying social conceptual development is also critical for successful false belief reasoning. However, the factors that are important for conceptual development are not addressed by this account; instead, focus is placed on the role that RC-EF skills play in
online false belief reasoning. This account is compatible with evidence for concurrent relations between RC-EF skills and false belief reasoning in both children and adults. However, it cannot easily explain a unidirectional, longitudinal relation between early RC-EF and later false belief, particularly when later, concurrent RC-EF skills are controlled for.

Finally, the emergence account takes a conceptual-developmental approach by stipulating that RC-EF skills are involved in the process of experience-driven social-cognitive theory change. From this account, RC-EF skills are related to false belief task performance because they facilitate the social conceptual developments that are necessary for successful false belief reasoning. Unlike the early understanding and component process accounts, this theory offers a straightforward explanation for the findings of longitudinal studies on the relation between RC-EF skills and false belief understanding. However, no studies to date have directly tested the extent that RC-EF skills interact with relevant experience over time to predict subsequent performance on false belief tasks.

**Goals of the Present Study**

While the emergence account is in line with much of the accrued research on the relation between RC-EF and false belief understanding, no studies to date have directly tested the central premise of this account. That is, no studies have examined whether RC-EF skills interact with everyday false-belief-relevant experience to contribute to naturalistic, protracted trajectories of false belief development over the preschool period. The goal of this project was to provide a powerful test of the emergence account by
examining whether RC-EF skills and natural forms of relevant experience interact to affect false belief abilities, and whether this pattern of findings exists most strongly when measured concurrently or longitudinally. I also aimed to test whether interactive effects involving early RC-EF skills exist independent of later RC-EF skills, and vice versa.

To test these questions, young preschool-aged children were recruited to participate in a longitudinal study over a 6-month period. During the initial testing phase, children were assessed on their false belief understanding and RC-EF skills. To index natural forms of false-belief-relevant experience, children’s parents were also asked to narrate a wordless storybook to their children. Parents’ narratives were then coded for mental state term references, thereby providing a well-established index of parent mental state talk (e.g., Ruffman et al., 2002). As a second measure of false-belief-relevant experience, data was collected on the number of child-aged siblings living in each participant’s home (e.g., Ruffman et al., 1998). In addition to these focal measures, control measures were administered to assess receptive vocabulary (Dunn & Dunn, 1997), and precursor theory-of-mind knowledge (Wellman & Liu, 2004). During the second phase of testing, which occurred 6 months following the first, families returned to the lab and all measures were re-administered.

As highlighted above, only the emergence account stipulates an interaction between RC-EF and relevant experience in the process of false belief conceptual change. Nevertheless, it is important to note that on the level of statistical relations among variables, the component process account is also compatible with finding an interaction between RC-
EF and relevant experience. To illustrate, suppose that RC-EF skills are important for online false belief reasoning via self-perspective inhibition, as the component process account would suggest. Let us further suppose that relevant experience is independently important for social conceptual development, which in turn is necessary for successful performance on false belief tasks (via its importance for other perspective taking). If both RC-EF skills and relevant experience are independently necessary but not sufficient for false belief reasoning, then a statistical interaction between variables might emerge whereby children with both advanced RC-EF skills and high levels of false-belief-relevant experience outperform their peers on false belief tasks. This statistical interaction would be possible, even if RC-EF skills and relevant experience do not interact on a mechanistic level in the process of conceptual change.

With this in mind, simply identifying a statistical interaction between RC-EF skills and relevant experience was not enough to discriminate between predictions made by the component process and emergence accounts. Instead, testing the emergence account’s predictions as separate from the component process account further necessitated testing the relative and independent strength of concurrent and longitudinal interactions among measures. In the following paragraphs I describe this study’s focal analyses, specifying concurrent and longitudinal predictions in turn.

**Concurrent analyses and predictions.** Focal analyses first tested for the interactive relation between RC-EF skills and relevant experience concurrently, at phase 1 and phase 2 independently. I hypothesized that there would be an interactive effect of RC-
EF skills and relevant experience on false belief understanding at both time points. More specifically, I expected that only children with advanced RC-EF skills would show an effect of relevant experience on false belief performance. Importantly, the emergence account would predict this concurrent pattern of relations to exist as a byproduct of the prolonged, longitudinal effects of RC-EF skills on experience-driven false belief conceptual development. With this in mind, I ran control analyses to test whether concurrent relations among measures at phase 2 remained significant when controlling for phase 1 RC-EF skills. I predicted that the interactive relation would drop below significance when the longitudinal effects of RC-EF skills were controlled in this way.

While the emergence account would not predict concurrent, interactive relations to exist independent of early RC-EF effects, this pattern of findings would be more straightforwardly compatible with the component process account. According to the component process account, RC-EF skills enable children to engage in self-perspective inhibition when faced with false belief reasoning scenarios. If RC-EF skills are necessary for online self-perspective inhibition as the concurrent process account asserts, and if relevant experience is independently important for false belief conceptual development, then we might expect to find children with high RC-EF skills and high levels of relevant experience to be uniquely successful on concurrently-measured false belief tasks. This effect would be expected to exist independent of the longitudinal effects of RC-EF. Taken together, this study’s concurrent analyses offered one means of testing between key predictions of the component process and emergence accounts (see Table 1 for a summary).
Table 1
A Summary of the Predictions made by the Component Process and Emergence Accounts.

<table>
<thead>
<tr>
<th></th>
<th>Concurrent Analysis Predictions</th>
<th>Longitudinal Analysis Predictions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Component Process Account</strong></td>
<td>- Compatible with finding an interactive effect of RC-EF and relevant experience on concurrent false belief performance.</td>
<td>- Compatible with finding an interactive effect of phase 1 RC-EF and relevant experience on phase 2 false belief understanding.</td>
</tr>
<tr>
<td></td>
<td>- The phase 2 concurrent interactive effect should exist independent of phase 1 RC-EF skills.</td>
<td>- The longitudinal interactive effect should not exist independent of phase 2 (concurrent) RC-EF skills.</td>
</tr>
</tbody>
</table>

| **Emergence Account** | - Compatible with finding an interactive effect of RC-EF and relevant experience on concurrent false belief understanding. | - Predicts an interactive effect of phase 1 RC-EF and relevant experience on phase 2 false belief understanding. |
|                        | - The phase 2 concurrent interactive effect should *not* exist independent of phase 1 RC-EF skills. | - This longitudinal interactive effect should exist independent of phase 2 (concurrent) RC-EF skills. |

**Longitudinal analyses and predictions.** Subsequent focal analyses examined the longitudinal interaction between early RC-EF skills and experience in predicting later false belief reasoning. In line with the emergence account, I hypothesized that early RC-EF skills...
and experience would interact to predict later false belief understanding, with relevant experience uniquely benefitting children with advanced RC-EF skills. Further, I predicted that these longitudinal interactive effects would exist independent of any concurrent effects of RC-EF skills on false belief performance. This pattern of results would be predicted by the emergence account, but not easily explained by the component process account of the relation between RC-EF and false belief. This study’s longitudinal analyses therefore offered a second means of testing between these theories (see Table 1).

Following from other longitudinal research interested in causal effects (e.g., Carlson et al., 2004; Hughes, 1998), I also tested the reverse interactive relation (i.e., the interaction between phase 2 RC-EF skills and experience in association with phase 1 false belief reasoning). I further predicted that the reverse developmental interaction—between later RC-EF skills and experience predicting early false belief understanding—would not be significant. This finding would align with prior research showing that while early RC-EF skills predict subsequent advances in false belief reasoning, the reverse developmental relation does not typically hold (e.g., Carlson et al., 2004; Hughes, 1998).

**The role of age, vocabulary skills, and initial theory-of-mind knowledge.**

Finally, I ran control analyses to test the effects of three relevant control variables on my concurrent and longitudinal findings. I predicted that any interactive relations among the focal constructs of interest would exist independent of these control variables. The first variable under consideration was children’s age. When considering conceptual change of any kind, one might expect older children to generally be in a better position to benefit
from experience in the service of advancing their understandings. Controlling for age provided a means of testing relations between variables independent of general, age-related maturation.

The second control variable was a measure of children’s receptive vocabulary skills (Dunn & Dunn, 1997). Many researchers have theorized a role for aspects of language abilities in the development of false belief understanding (see Astington & Baird, 2005). Indeed, research shows a robust relation between children’s receptive vocabulary and both false belief and RC-EF task performance (e.g., Carlson & Moses, 2001). Including a receptive vocabulary task in this study provided a means of both testing and controlling for the role that vocabulary abilities play in false belief development over time.

The final control measure was a theory-of-mind scale that assessed children’s understanding of the mental state concepts that typically come to be understood just prior to understanding false beliefs (i.e., diverse desires, diverse beliefs, and knowledge vs. ignorance) (Wellman & Liu, 2004). Many researchers have theorized the importance of prior knowledge for both constraining and driving subsequent conceptual development (e.g., Gopnik et al., 2004). In line with this notion, one recent study demonstrated that initial theory-of-mind knowledge was a significant predictor of subsequent advances in false belief reasoning (Rhodes & Wellman, 2013). Including the theory-of-mind scale in this study provided a means of accounting for the contributions made by precursor theory-of-mind knowledge to false belief task performance across the testing period.
Chapter 2

Method

Participants

Phase 1. Participants were 81 preschool-aged children (36 males, 45 females). At the time of the initial testing session, children ranged in age from 3 years, 5 months to 4 years, 0 months ($M = 3$ years, 8 months, $s = 1.90$ months). Eleven additional children were recruited but excluded from the study’s analyses because of failure to return to the lab for the second study session ($n = 5$), technical difficulties leading to the loss of session data ($n = 2$), the child’s refusal to complete one or more of the behavioural tasks administered ($n = 2$), or the child obtaining a score that was more than 3 standard deviations beyond our sample’s mean on one of our key variables of interest ($n = 2$). Participants were recruited through local-area daycares, and through a database that was formed based on information from birth announcements and from parents who signed up at fairs, festivals, and other events in the area. The sample was representative of the primarily white, middle-class community in Southeastern Ontario from which they were recruited.

Phase 2. In order to measure changes in false belief understanding over time, a selection of participants were invited to return to the lab for a second phase of testing 6 months following the first. Of the 81 children included in phase 1, 55 demonstrated poor performance on the initial false belief battery (i.e., scores of 50% or less) and were thus invited to return for phase 2. The remaining children were excluded from phase 2 participation because they had already established an understanding of false beliefs at phase
and therefore had limited room to improve on the false belief battery over time. The final phase 2 sample consisted of 41 children (20 male, 21 female), who ranged in age from 3 years 11 months to 4 years 7 months ($M = 4$ years, 2 months, $s = 1.93$ months). Two additional children participated in the phase 2 sessions but were excluded from analyses because they obtained a score that was more than 3 standard deviations beyond the mean on one of our key variables of interest.

**Materials and Measures**

All measures were administered twice, during both phase 1 and 2 of this study. Focal measures were a false belief battery, a set of executive functioning tasks, and 2 indices of false-belief-relevant experience. To examine alternative explanations for our study findings, 2 additional measures were administered assessing children’s receptive vocabulary abilities, and their more general theory-of-mind knowledge. Each of these measures is described in more detail below. To examine questions that extend beyond the scope of this dissertation, I also obtained a measure of children’s resting EEG, and administered a questionnaire aimed at collecting information on pets living in participants’ homes.

**False Belief battery.** Sample scripts for the 4 tasks included in the false belief battery are presented in Appendix A. The tasks included in the battery measure children’s ability to identify their own and others’ false beliefs, and predict resulting behaviour. All tasks were presented to the children using toy figurines, puppets and picture props. A brief description of the tasks included in this battery follows:
**Contents Change task** (Gopnik & Astington, 1988; Perner, Leekham, & Wimmer, 1987). Children were shown a container that would normally contain something familiar (e.g., a Smarties box) and asked to state what it contains. The experimenter went on to show them that the box actually contained something unexpected (e.g., pencils). Children were introduced to a character who had never seen inside the box before. They were then asked the memory control question: “Did [character’s name] see inside the box?” If children answered incorrectly (Phase 1: \( n = 21 \), Phase 2: \( n = 2 \)), they were corrected and asked the memory control question again until they answered correctly. Children were then asked the test question, e.g., “What does [character’s name] think is in the box, Smarties or pencils?” Children were given a score of 1 if they answered the test question correctly.

**Location Change task** (Wimmer & Perner, 1983). A character was shown to hide an object in one location and leave the scene. In her absence, a second character moved the object from its original location and hid it somewhere else. The first character then returned, and children were asked the control question: “Did [character’s name] see the [object] get moved?” If children responded incorrectly (Phase 1: \( n = 27 \), Phase 2: \( n = 3 \)), the story was re-enacted and the control question was posed again. After children responded correctly to the control question, the test question was asked: “When [character’s name] comes back, where will he look for the [object]?” Children were given a score of 1 if they correctly answered the test question.

**Appearance-Reality task** (Flavell, Green, & Flavell, 1986). This task included 2 trials. In the first trial, children were shown an object that looked like one thing, but was
actually another (e.g., a sponge that was decorated to look like a rock). Children were presented with two test questions. The first asked how the object looked (e.g., “When you look at this right now, does it look like a sponge, or a rock?”). The second asked what the object actually was (e.g., “What is this really and truly, a sponge or a rock?”). The second trial involved showing children a picture of an object that appeared to be a different colour when covered by colored film (e.g., an orange castle that looked black under a green film). The film was placed over the picture, and children were asked what colour the object looked (e.g., “When you look at this castle right now, does it look orange or black?”), and what colour the object was in reality (e.g., “What colour is the castle, really and truly?”).

On each trial, children received a score of 1 if they correctly responded to both test questions (0-2 possible).

**Deceptive Pointing task** (Carlson et al., 1998). Children saw an experimenter place a toy into one of two closed boxes. As the experimenter pointed to the box, she explained to the children that you can indicate an object’s whereabouts by pointing to where it is hidden. After practicing using a point gesture to show which of two boxes an object was placed in, children were encouraged to play a trick on a puppet to make her look in an empty box for the hidden toy. Children participated in 2 trials that each involved playing a trick on a different character. On each trial, children were awarded a score of 1 if they successfully pointed to the empty box (0-2 possible).
Executive Functioning Tasks. These tasks require children to inhibit their initial impulses in order to respond in alternate ways. The complete scripts and scoring protocol are given in Appendix B.

Bear/Dragon task (Kochanska, Murray, Jacques, & Koenig, 1996; Reed, Pein, & Rothbart, 1984). This task is a modified version of a "Simon Says" game. The experimenter first asked children to perform 10 simple, clearly defined actions (e.g., touch your head). They were then introduced to a bear puppet and a dragon puppet, and instructed to do what they were told by the "nice bear", but not what they were told by the "mean dragon". The puppets, which were both controlled by the experimenter, then took turns telling the child to act out the actions that were performed earlier in the game. Practice trials where children were given feedback on their answers were administered until the children correctly responded to one bear trial and one dragon trial. Responses on the dragon test trials were coded as follows: 0 = full correct movement, 1 = partial correct movement, 2 = incorrect movement (e.g., touches feet when instructed to touch head), 3 = strategic movement (e.g., shakes head no, moves hands behind back), and 4 = no movement. Standardized scores were calculated based on children’s average performance across trials. Inter-rater reliability was calculated for 30% of participants and Kappa was .93 at phase 1 and .96 at phase 2, with an overall accuracy of 96.00% at phase 1 and 98.39% at phase 2.

Grass/Snow task (Carlson & Moses, 2001). A black board with a white card attached to the upper left-hand corner and a green card attached to the upper right-hand corner was placed in front of the children. Children were asked to play a silly game by
pointing to the green square anytime the experimenter said the word “snow”, and to the white square anytime she said the word “grass”. Practice trials where children were given feedback on their answers were administered until the children correctly responded to one grass trial and one snow trial consecutively. Following this, there were 16 test trials without feedback. On each of the test trials, children were given a score of 1 for initially pointing to the correct square and 0 for any other response. Inter-rater reliability was calculated for 30% of participants and Kappa was .97 at phase 1 and .95 at phase 2, with an overall accuracy of 98.50% at phase 1 and 98.56% at phase 2.

**Dimensional Change Card Sort (DCCS) task** (Frye, Zelazo, & Palfai, 1995; Zelazo, Müller, Frye, & Marcovitch, 2003). Children were given a deck of cards. Each card depicted one of two shapes (boat or rabbit) in one of two colors (red or blue). First, children were asked to sort the cards into two piles based on shape (i.e., boats cards go in one pile, rabbit cards in another). Then, children were asked to change rules and sort the remaining five cards into the same locations based on their colour (i.e., red cards go in one pile, blue cards in another). Children were given a score out of 5 based on the number of post-switch cards that were correctly sorted.

**Measures of Relevant Experience.** I measured two types of experiences that prior research has shown correlate with false belief development: parent mental state talk (e.g., Ruffman et al., 2002), and number of child-aged siblings in the home (e.g., Perner et al., 1994; Ruffman et al., 1998).
**Parent mental state talk.** As a measure of parent mental state talk, parents and children participated in a well-established picture-book reading task. Parents and children were brought into a quiet room and provided with an illustrated, wordless storybook entitled *The Snowman* (Briggs, 1978) during phase 1, and *The Midnight Circus* (Collington, 1992) during phase 2. Each story depicted the adventures of a child character. For instance, *The Snowman* depicts the adventures of a young boy who builds a snowman that magically comes to life. During each phase of the study, parents were asked to narrate the story to their children as they would a story at home. The narratives provided by parents were video-recorded and subsequently transcribed and coded both for number of total words uttered, and for the extent that parents made reference to mental states throughout the story.

Varying the story from phase 1 to phase 2 ensured that both children’s and parents’ responses to the story during the second phase of testing were not influenced by having previously read the same story. It also provided two snapshots of parent mental state talk in different narrative contexts, which themselves offered varied opportunities to reference other minds. It is important to note that parent mental state talk while narrating stories generally—or these books in particular—is not necessarily uniquely meaningful for false belief development, over and above parent mental state talk in other types of interactions. Nevertheless, storybook tasks such as this have proven to be a particularly fruitful means of capturing parents’ tendency to focus on mental states when communicating with their children (Sabbagh & Callanan, 1998). It is also noteworthy that this study paradigm was not designed to measure meaningful changes in parent mental state talk across time;
instead, our goal was to obtain as robust a measure of mental state talk as possible by
quantifying the same construct on two separate occasions in different narrative contexts.

Using mental state term lists provided by Ruffman and colleagues as a guide
(Ruffman et al., 2002), two coders independently combed each dyad’s transcripts and
identified all clear instances of parent-uttered mental state terms. During this initial coding
phase, each coder also flagged other utterances that she thought had potential to be coded as
a mental state term (e.g., “freaked out”, which was ultimately deemed to be a synonym for
“scared” and coded as an emotion term). From this work, an exhaustive list of all potential
mental state utterances was created for each dyad.

All utterances identified through this initial process were then categorized by a
single coder using criteria outlined by Ruffman et al. (2002). Table 2 provides a list of
mental state categories and sample words coded. The following 4 general coding rules
applied across categories:

1) When a particular mental state term was uttered more than once with no more
than 2 sentences between instances, subsequent utterances were only coded as
additional mental state terms if the subject of the mental state changed. For
instance, when coding the sentence, “Do you think that’s tea or do you think
that’s juice?” the mental state term “think” was counted once. I reasoned that
these instances only warrant one coding because the same meaning could easily
be conveyed by saying “Do you think that’s tea or juice?” Indeed, many parents
tended to produce these types of statements using the latter, more concise format.
In contrast, when coding “I didn’t know snowmen could fly. Did you know snowmen could fly?” the mental state term “know” was counted twice because the subject of the mental state term changed from one instance to the next.

2) A mental state term uttered by the child and subsequently repeated by the parent was not coded as such (e.g., Child stating “I like the red ones” followed immediately by the parent’s stating, “Yeah, you like the red ones”).

3) A mental state term within a sentence fragment (i.e., a sentence that was abandoned mid-phrase, before its meaning was conveyed) was not counted as such (e.g., “He wants to—he’s looking at his snowman”).

4) Mental state terms uttered as a part of common sayings were excluded from coding, (e.g. “What do you know?”, “pleased to meet you”, “the show must go on”).
Table 2

*Categories and Examples of Mental State Terms*

<table>
<thead>
<tr>
<th>Category</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge</td>
<td>Think, know, believe</td>
</tr>
<tr>
<td>Modulation of assertion</td>
<td>Maybe, might, probably, possibly, guess, sure, suppose, perhaps, could be, must, seems, certainly, definitely, figure, suppose, bet, actually, right, that’s right, you’re right, really, or something</td>
</tr>
<tr>
<td>Desire</td>
<td>Want, like, love, wish, prefer, hope, favourite, rather, enjoy</td>
</tr>
<tr>
<td>Emotion</td>
<td>Happy, sad, unhappy, angry, excited, upset, scared, surprised, miss, worry, pleased, feel badly, glad, intrigued, interested, relieved, delighted, feel better, frightened, thankful, disappointed, engrossing, mad, afraid, horrified, impressed, shocked, ticked off, freaked out, grateful, proud, satisfied, defeated</td>
</tr>
<tr>
<td>Pretend</td>
<td>Pretend, imagine</td>
</tr>
<tr>
<td>Other</td>
<td>Dream, wonder, remember, understand, tricked, pay attention, remind, find out (when object was specified), figure out (when object was specified), figure, expect, confused</td>
</tr>
</tbody>
</table>

Additional category-specific coding rules are outlined in the following paragraphs:

*Knowledge.* Instances of “think”, “know”, “believe”, and their derivatives were coded as knowledge terms, except for 2 instances: 1) When parents used the phrase “you know?” or “I know!” without an object following the verb “know”, and 2) when parents uttered the 3-word, complete sentence “I don’t know”. As argued by Ruffman et al. (2002),
“I don’t know” could easily be interpreted simply as “I don’t have an answer for you”, rather than being a more direct reference to a mental state.

Modulation of Assertion. Modulation of assertion terms were words deemed to provide information on the certainty of a statement or notion, including “might”, “maybe”, “must”, and “definitely” (see Table 2). “Think” and “know” are also terms often used to modify certainty, but following from previous work (e.g., Ruffman et al., 2002), they were coded separately as knowledge terms, rather than modulations of assertion terms. When parents used more than one modulation of assertion term in reference to a single statement, only the first instance was counted.

Desire: Several words and their derivatives were coded as references to desires, including “like”, “love”, “want”, and “admire” (see Table 2). The only exception to this coding came in instances where a desire term was used to convey a command (e.g., “I’d like you to sit beside me, please.” or “Want to put that wrapper in the garbage, please?”)

Emotion. Emotion terms were words that made reference to feelings, including happy, sad, and surprised (see Table 2). The term “feel” was only coded as an emotion term when followed by a word that clarified it’s intent as a true emotion reference (e.g., “feel sad” or “feel bad”), and not when the term was uttered on its own (e.g., “how does she feel?”). Following from Ruffman et al. (2002), I reasoned that when “feel” was uttered on its own, it may be referring to a physical state (e.g., “he feels cold”), rather than a mental state. Also excluded were emotion terms applied to non-agents (e.g., “This story is scary”), as they did not reference a specific character or person’s mental state.
Pretend. Pretend terms included the words “pretend”, “imagine” and their derivatives. All instances of these terms were coded.

Other. The “other” mental state category included mental state terms that did not fit into any of the categories listed above, including “wonder”, “remember”, and “dream” (See Table 2).

Words that were flagged in the initial coding phase but that subsequently did not meet criteria for inclusion in any of the mental state categories were coded as non-mental utterances. Upon closer examination of parents’ mental state talk after the initial coding phase, the decision was made to code three words as non-mental in nature: decide, need, and learn. These words were not mentioned in prior mental state talk research (e.g., Ruffman et al., 2002), but were originally flagged by coders as potential mental state utterances. However, careful examination of the use of these terms revealed that they often seemed not to refer to true mental functions. For instance, while the term ‘need’ at times may have been used interchangeably with the term ‘want’ (e.g., “I need a scarf for my snowman”), the message conveyed tended to focus on concrete physical requirements, rather than subjective desires. Similarly, while the term ‘learn’ has the potential to be used as a reference to conceptual knowledge acquisition, it seemed most commonly to be used as reference to new physical abilities (e.g., “learn how to turn the lights on and off”).

To obtain a measure of reliability, a second coder independently coded 30% of participants’ word lists using the criteria outlined above. Kappa was a very respectable .99
at phase 1 and .98 at phase 2, with an overall accuracy of 98.90% at phase 1 and 98.75% at phase 2. All discrepancies in coding were identified and resolved through discussion.

**Number of child-aged siblings in the home.** Number of child-aged siblings was our second measure of relevant experience. At both phases, parents were asked to list the number and ages of their child’s siblings in the home (See Appendix C). Following from prior research showing that child-aged siblings uniquely impact false belief reasoning (e.g., Ruffman et al., 1998), sibling scores were calculated by tallying the number of siblings aged 4 to 8 living with each of our child participants.

**Vocabulary Measure.** The Peabody Picture Vocabulary Test, Fourth Edition was used to test children’s receptive vocabulary (PPVT-IV, Dunn & Dunn, 1997). In each trial, children were shown a page with four drawings, and asked to point to the drawing that corresponded with a verbally-presented vocabulary word (e.g., “Point to dog.”). No feedback was provided on children’s answers to test trials. Children continued through progressively more difficult items until they incorrectly answered more than 8 in a block of 12. Children’s scores on the measure were calculated by subtracting the total number of errors made from the highest item reached, as outlined in the instruction manual (Dunn & Dunn, 1997). The same version of the PPVT-IV was used at both time points.

**Theory of Mind Scale** (Wellman & Liu, 2004). Scripts for each task included are presented in Appendix D. The tasks administered assess children’s understanding of mental state concepts that typically come to be understood prior to understanding false beliefs (i.e., diverse desires, diverse beliefs, and knowledge vs. ignorance, Wellman & Liu, 2004). Toy
figurines, puppets and picture props were used to illustrate the task scenarios. The tasks were administered in a set order, becoming progressively more difficult. A brief description of each task included in the scale is provided below.

**Diverse Desires task.** Children were shown two types of snacks, and asked to state which snack they would prefer to eat. They were then informed that a character (Big Bird) prefers the alternative snack. Children were asked to state which snack Big Bird would choose to eat. In order to answer the question correctly, children had to understand that Big Bird would choose the snack that was in accordance with his own preferences, not their own.

**Diverse Beliefs task.** Children were shown a puppet (Cookie Monster), and presented with pictures of both a garage and a tree. They were told that Cookie Monster wants to find his cat, and that his cat could be hiding in either of the illustrated locations. Children were then asked to state where they thought the cat was. The experimenter proceeded to inform them that Cookie Monster thinks the cat is in the alternate location. Children were then asked to state where Cookie Monster will look for his cat. In order to answer the question correctly, children had to understand that Cookie Monster would look in the location that was in accordance with his beliefs, not their own.

**Knowledge Access task.** Children were shown a jewelry box and asked what they thought was inside. After they responded with an answer or stated that they did not know, the box was opened to reveal a small toy frog. The jewelry box was closed again, and children were introduced to a character (Samantha). They were told that Samantha has
never seen inside the box, and then asked whether Samantha knows its contents. In order to answer the question correctly, children had to understand that Samantha did not have the same knowledge about the box’s contents that they themselves had.

**Procedure**

This study involved 4 videotaped sessions, which were each between 30 and 60 minutes in length and took place in a university laboratory. The same female experimenter administered all study sessions.

**Phase 1.** Phase 1 was comprised of 2 sessions, which took place within 2 weeks of one another. At the beginning of the first session, children’s resting EEG was measured using a procedure outlined by Sabbagh, Bowman, Evraire, and Ito (2009). Following EEG data collection, both child and parent were led to a quiet room where the parent was instructed to narrate a wordless picture book to his or her child as they would a story at home. Finally, the experimenter sat with the child and administered the theory-of-mind scale. During the second study session, the experimenter administered the false belief battery, the RC-EF tasks, and the PPVT-IV. While the child was engaged with the experimenter, parents were asked to fill out questionnaires to provide information on child-aged siblings and family pets. Task order was kept consistent across participants to ensure that individual differences in children’s performance could not be attributed to variance in task administration.

**Phase 2.** Phase 2 occurred 6 months after phase 1, and involved repeating the phase 1 games and activities across 2 sessions. Again, the sessions took place within 2
weeks of one another. Across phases the surface characteristics of the tasks (e.g., objects and character names) were changed, but the underlying structure of each task and the order in which the tasks were administered within each session remained the same (See Table 3).
Table 3

Task Administration

<table>
<thead>
<tr>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase 1</td>
<td>Phase 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td>EE G</td>
<td>EEG Measurement</td>
<td>FB Contents Change</td>
<td>FB Contents Change</td>
</tr>
<tr>
<td>Parent Mental State</td>
<td>Parent Mental State</td>
<td>RC-EF Bear/Dragon</td>
<td>RC-EF Bear/Dragon</td>
</tr>
<tr>
<td>Talk Measure</td>
<td>FB Location Change</td>
<td>Talk Measure</td>
<td>FB Location Change</td>
</tr>
<tr>
<td>(Storybook Task)</td>
<td>RC-EF Grass/Snow</td>
<td>(Storybook Task)</td>
<td>RC-EF Grass/Snow</td>
</tr>
<tr>
<td>Theory of Mind Scale</td>
<td>FB Appearance/Reality</td>
<td>Theory of Mind Scale</td>
<td>FB Appearance/Reality</td>
</tr>
<tr>
<td>RC-EF Card Sort (DCCS)</td>
<td>FB Deceptive Pointing</td>
<td>RC-EF Card Sort (DCCS)</td>
<td>FB Deceptive Pointing</td>
</tr>
<tr>
<td>PPVT</td>
<td>Parent Questionnaires:</td>
<td>PPVT</td>
<td>Family Pets</td>
</tr>
<tr>
<td>Child-Aged Siblings</td>
<td>Child-Aged Siblings</td>
<td>Family Pets</td>
<td></td>
</tr>
<tr>
<td>Note. FB = False Belief; PPVT = Peabody Picture Vocabulary Test; DCCS = Dimensional Change Card Sort</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Chapter 3

Results

Performance on Measures Independently

Tables 4 and 5 provide a summary of descriptive statistics for our key measures. Note that statistics are reported both for our full sample at phase 1, and for the subset of children who returned for the second phase of testing, deemed the ‘returning subset’.

False Belief Battery. Performance on the 4 tasks comprising the false belief battery is reported in Tables 4 and 5. Scores at phase 1 were in line with prior research employing this battery with children in our age range (e.g., Carlson & Moses, 2001). Recall that false belief battery performance at phase 1 was used as the inclusion criteria for phase 2 of the study; only children who scored 50% or less on the initial false belief battery were included in the returning subset. Within the returning subset, performance on the false belief battery as a whole improved significantly from phase 1 to phase 2 (see Table 5).
Table 4

Descriptive Statistics of Measures within Full Sample at Phase 1 (n = 81)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Full Sample Phase 1 (n = 81)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (s)</td>
</tr>
<tr>
<td>False Belief Battery (/4)</td>
<td>1.69 (1.24)</td>
</tr>
<tr>
<td>Contents Change (pass/fail)</td>
<td>20 (24.69)</td>
</tr>
<tr>
<td>Location Change (pass/fail)</td>
<td>27 (33.33)</td>
</tr>
<tr>
<td>Appearance Reality (/2)</td>
<td>0.86 (0.82)</td>
</tr>
<tr>
<td>Deceptive Pointing (/2)</td>
<td>1.35 (0.90)</td>
</tr>
<tr>
<td>RC-EF Battery</td>
<td></td>
</tr>
<tr>
<td>Bear/Dragon Practice</td>
<td>2.11 (1.88)</td>
</tr>
<tr>
<td>Bear/Dragon Test (/20)</td>
<td>13.42 (7.94)</td>
</tr>
<tr>
<td>Grass/Snow Practice</td>
<td>2.85 (1.47)</td>
</tr>
<tr>
<td>Grass/Snow Test (/16)</td>
<td>10.17 (4.63)</td>
</tr>
<tr>
<td>Card Sort (/5)</td>
<td>3.46 (1.42)</td>
</tr>
<tr>
<td>Measures of Experience</td>
<td></td>
</tr>
<tr>
<td>Child-aged Siblings</td>
<td>0.52 (0.59)</td>
</tr>
<tr>
<td>Mental State Talk (/1000 words)</td>
<td></td>
</tr>
<tr>
<td>Knowledge</td>
<td>7.73 (5.15)</td>
</tr>
<tr>
<td>Modulation of Assertion</td>
<td>3.90 (2.86)</td>
</tr>
<tr>
<td>Desire</td>
<td>4.87 (2.47)</td>
</tr>
<tr>
<td>Emotion</td>
<td>1.89 (1.81)</td>
</tr>
<tr>
<td>Pretense</td>
<td>0.26 (0.43)</td>
</tr>
<tr>
<td>Other</td>
<td>1.22 (1.50)</td>
</tr>
<tr>
<td>Total Narrative Word Count</td>
<td>1251.42 (386.04)</td>
</tr>
<tr>
<td>PPVT-IV</td>
<td>80.91 (15.68)</td>
</tr>
<tr>
<td>Theory-of-Mind Scale (/3)</td>
<td>1.84 (0.99)</td>
</tr>
<tr>
<td>Diverse Desires (pass/fail)</td>
<td>58 (71.60)</td>
</tr>
<tr>
<td>Diverse Beliefs (pass/fail)</td>
<td>54 (66.66)</td>
</tr>
<tr>
<td>Knowledge Access (pass/fail)</td>
<td>36 (44.44)</td>
</tr>
</tbody>
</table>

Note. *Descriptive statistics reported for dichotomous variables are: n correct (% correct).
Table 5
Descriptive Statistics of Measures within Returning Subset at Phases 1 and 2 (n = 41)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Phase 1: M (s)</th>
<th>Phase 2: M (s)</th>
<th>Difference Between Phases</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>False Belief Battery (/4)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contents Change (pass/fail)</td>
<td>2 (4.87)</td>
<td>13 (31.71)</td>
<td>$t(40) = 8.42, p &lt; .001, d = 1.49$</td>
</tr>
<tr>
<td>Location Change (pass/fail)</td>
<td>3 (7.31)</td>
<td>18 (43.90)</td>
<td>$X^2 (1) = 13.07, p &lt; .001, \phi = .56$</td>
</tr>
<tr>
<td>Appearance Reality (/2)</td>
<td>0.73 (0.74)</td>
<td>1.15 (0.85)</td>
<td>$t(40) = 2.88, p = .006, d = 0.45$</td>
</tr>
<tr>
<td>Deceptive Pointing (/2)</td>
<td>1.15 (0.94)</td>
<td>1.83 (0.50)</td>
<td>$t(40) = 4.44, p &lt; .001, d = 0.74$</td>
</tr>
<tr>
<td><strong>RC-EF Battery</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bear/Dragon Practice</td>
<td>2.61 (2.08)</td>
<td>1.07 (0.35)</td>
<td>$t(40) = -4.77, p &lt; .001, d = 0.97$</td>
</tr>
<tr>
<td>Bear/Dragon Test (/20)</td>
<td>11.51 (8.06)</td>
<td>18.17 (2.54)</td>
<td>$t(40) = 5.83, p &lt; .001, d = 1.18$</td>
</tr>
<tr>
<td>Grass/Snow Practice</td>
<td>3.24 (1.79)</td>
<td>2.29 (0.64)</td>
<td>$t(40) = -3.46, p = .001, d = 0.63$</td>
</tr>
<tr>
<td>Grass/Snow Test (/16)</td>
<td>9.66 (4.67)</td>
<td>12.54 (3.12)</td>
<td>$t(40) = 4.14, p &lt; .001, d = 0.68$</td>
</tr>
<tr>
<td>Card Sort (/5)</td>
<td>3.49 (1.49)</td>
<td>4.15 (1.30)</td>
<td>$t(40) = 2.60, p = .013, d = 0.41$</td>
</tr>
<tr>
<td><strong>Measures of Experience</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child-aged Siblings</td>
<td>0.51 (0.60)</td>
<td>0.49 (0.55)</td>
<td>n/a</td>
</tr>
<tr>
<td><strong>Mental State Talk (/1000 words)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knowledge</td>
<td>6.76 (4.30)</td>
<td>5.95 (4.47)</td>
<td>$t(40) = 1.42, p = .162$</td>
</tr>
<tr>
<td>Modulation of Assertion</td>
<td>3.72 (2.71)</td>
<td>3.49 (2.61)</td>
<td>$t(40) = 0.532, p = .597$</td>
</tr>
<tr>
<td>Desire</td>
<td>5.00 (2.48)</td>
<td>3.94 (2.04)</td>
<td>$t(40) = 2.37, p = .023, d = .985$</td>
</tr>
<tr>
<td>Emotion</td>
<td>1.77 (2.12)</td>
<td>7.28 (2.94)</td>
<td>$t(40) = 12.20, p &lt; .001 d = 1.85$</td>
</tr>
<tr>
<td>Pretense</td>
<td>0.23 (0.37)</td>
<td>0.22 (0.46)</td>
<td>$t(40) = 0.05, p = .962$</td>
</tr>
<tr>
<td>Other</td>
<td>1.17 (1.50)</td>
<td>1.73 (1.76)</td>
<td>$t(40) = 2.06, p = .046, d = .98$</td>
</tr>
<tr>
<td>Total Narrative Word Count</td>
<td>1314.71 (414.25)</td>
<td>1201.00 (391.87)</td>
<td>$t(40) = 2.02, p = .051$</td>
</tr>
<tr>
<td>PPVT-IV</td>
<td>77.27 (12.76)</td>
<td>89.71 (14.76)</td>
<td>$t(40) = 6.24, p &lt; .001, d = 1.01$</td>
</tr>
<tr>
<td>Theory-of-Mind Scale (/3)</td>
<td>1.56 (1.00)</td>
<td>2.41 (0.87)</td>
<td>$t(40) = 4.84, p &lt; .001, d = 0.77$</td>
</tr>
<tr>
<td>Diverse Desires (pass/fail)</td>
<td>26 (63.41)</td>
<td>36 (87.80)</td>
<td>$X^2 (1) = 5.79, p = .016, \phi = .38$</td>
</tr>
<tr>
<td>Diverse Beliefs (pass/fail)</td>
<td>26 (63.41)</td>
<td>31 (75.61)</td>
<td>$X^2 (1) = 1.07, p = .302, \phi = .16$</td>
</tr>
<tr>
<td>Knowledge Access (pass/fail)</td>
<td>12 (29.27)</td>
<td>32 (78.05)</td>
<td>$X^2 (1) = 18.05, p &lt; .001, \phi = .66$</td>
</tr>
</tbody>
</table>

Note. a Descriptive statistics reported for dichotomous variables are: n correct (% correct).
**RC-EF Battery.**

**Bear/Dragon task.** Both administrations of the Bear/Dragon task involved 2 dependent measures: number of attempts taken to pass the practice trials and score on the test trials (see Tables 4 and 5). Within the returning subset, performance on both practice and test trials improved from phase 1 to phase 2 (see Table 5). Following standard procedure, an overall phase 1 Bear/Dragon aggregate was calculated for both the full sample and returning subset by combining the standardized scores on practice trials (reverse scored) and test trials. At phase 2, all but 2 children performed perfectly on the Bear/Dragon practice trials, likely because they were already familiar with the rules of the game after having participated in phase 1 of the study. Thus, the phase 2 Bear/Dragon scores consisted of standardized scores on test trials, with performance on practice trials omitted.

**Grass/Snow task.** Both administrations of Grass/Snow involved 2 dependent measures: number of attempts taken to pass the practice trials and scores on the test trials (see Tables 4 and 5). Within the returning subset, performance on both practice and test trials improved from phase 1 to phase 2 (see Table 5). Following standard procedure, aggregate phase 1 and 2 Grass/Snow scores were calculated by combining the standardized dependent measures at each time point.

**DCCS task.** Children were scored on five post-switch trials where they were asked to sort cards according to colour after previously having to sort by shape (see Tables 4 and 5). Within the returning subset, performance on the post-switch trials improved from phase 1 to phase 2 (see Table 5).
Relations among Tasks and Scale Aggregating. At phase 1, inter-correlations among Bear/Dragon aggregate, Grass/Snow aggregate, and DCCS were all significant, and in line with prior research (full sample: $rs (79) = .35$ to $.72$, $ps < .05$, returning subset: $rs (39) = .32$ to $.64$, $ps < .05$). At phase 2, the inter-correlations among tasks were positive, though somewhat lower than at phase 1 ($rs (39) = .26$ to $.48$, $ps = .002$ to $.103$). Following prior research, an RC-EF aggregate was computed for each phase of the study by combining standardized scores on each of the three RC-EF measures.

Both RC-EF aggregate histograms and Q-Q plots were used to examine the shape of the RC-EF aggregate distributions. Q-Q plots (also referred to as normal probability plots) offer a means of graphically representing the extent that scores on a measure fit a normal distribution (De Veaux, Velleman, & Bock, 2006). If the distribution under examination is approximately normal, then the Q-Q plot will follow the $y = x$ line, while curves in the plot suggest a skewed distribution. Results of these tests indicated that RC-EF aggregate scores were negatively skewed, with a larger-than-expected proportion of our sample obtaining $z$-scores in the highest range of our distribution (see Figures 2 to 4). This was a consequence of a large proportion of children performing similarly and well overall on the RC-EF tasks, resulting in little meaningful variability in the upper range of our distribution. With this in mind, I opted to dichotomize our RC-EF battery scores using a mean split. Dichotomizing using a mean split ensured that the cluster of children who were performing similarly to one another at the upper range of our distribution were categorized together in the high RC-EF group, with the spread of children who obtained lower scores on our measure comprising the low RC-EF group.
Figure 2. Histogram and Q-Q Plot Illustrating RC-EF Aggregate Scores in the Full Sample at Phase 1 (n = 81)
Figure 3. Histogram and Q-Q Plot Illustrating RC-EF Aggregate Scores in the Returning Subset at Phase 1 (n = 41)
Figure 4. Histogram and Q-Q Plot Illustrating RC-EF Aggregate Scores in the Returning Subset at Phase 2 (n = 41)
Measures of Relevant Experience.

Parent mental state talk. Research suggests that relations between false belief understanding and parental mental state talk exist independent of story length (e.g., Ruffman et al., 2002). Thus, mental state talk frequency scores were created for each category by calculating the total number of mental state words uttered per 1000 words by each parent during their stories (see Tables 4 and 5). Recall that the story upon which parents’ narrations were based changed from phase 1 to phase 2. There were 2 instances where rates of mental state references changed across stories: parents tended to refer to desires significantly less and emotions significantly more when reading The Midnight Circus as compared to when reading The Snowman (see Table 5).

While many categories of mental state talk were coded, prior research has highlighted that two categories – knowledge and modulation of assertion—are most commonly related to false belief understanding (e.g., Ruffman et al., 2002). A mental state aggregate score for each phase was computed by combining standardized frequency scores for both the knowledge and modulation of assertion categories. Individual differences in mental state aggregate scores were highly correlated across phases, \( r (39) = .62, p < .001, \) and there were no significant differences in rates of these mental state utterances during The Snowman as compared to The Midnight Circus (see Table 5). I reasoned that measuring parent mental state talk on two occasions using two different stories was the best means possible of obtaining a robust measure of our construct of interest. Thus, an overall mental state aggregate score was computed for each parent by combining the standardized phase 1 and phase 2 mental state scores.
Child-aged siblings. Mean and standard deviations for child-aged siblings are reported in Tables 3 and 4. No child had more than 2 child-aged siblings. In the returning subset, child-aged sibling scores at phase 1 were identical to those at phase 2 in all but one case; one participant had a sibling who aged out of the sibling range in between phases. In the returning subset, sibling scores across the two phases were averaged to form an aggregate child-aged sibling score.

Experience aggregate. Although both child-aged siblings and mental state talk are thought to contribute to false belief development, I did not expect these two independent sources of relevant experience to be significantly related to one another. However, the correlation was significant in the full sample at phase 1, \( r(79) = .30, p = .006 \), indicating that children whose parents used more mental state terms in their stories also tended to have more child-aged siblings. It is unclear as to why this relation exists in our sample. Within the returning subset, mental state talk aggregate scores and child-aged sibling scores were not significantly related at either phase (Phase 1: \( r(39) = .17, p = .295 \); Phase 2: \( r(39) = -.10, p = .524 \)).

For study purposes, 2 experience aggregate scores were computed. The first was computed for use in analyses involving my full sample, and calculated by standardizing and combining phase 1 child-aged sibling scores and phase 1 mental state talk aggregate scores. The second was calculated for participants in our returning subset by combining standardized, aggregate scores on our sibling and mental state measures. Given that this aggregate included 2 measurements of each construct of interest (measurements from phases 1 and 2), I reasoned that it was the best possible estimate of each child’s exposure to
relevant experience. It was therefore employed as our measure of experience whenever possible (i.e., in all analyses conducted on children in the returning subset).

Inspection of experience aggregate histograms and Q-Q plots indicated that the distribution of experience aggregate scores computed on the full sample was bimodal (see Figures 5). This bimodality was likely driven by the inclusion of child-aged sibling scores in the aggregate, as most children received scores of 0 or 1 on the sibling measure. The returning subset’s experience aggregate histogram and Q-Q plot revealed a similar (albeit less pronounced) pattern to the full sample’s distribution (see Figure 6). In line with the RC-EF aggregate, I opted to dichotomize our experience aggregate scores (mean split). Dichotomizing allowed us to circumvent issues associated with attempting to establish linear relations between variables with atypically-shaped distributions.

**Vocabulary Measure.** Scores on the PPVT-IV at each time point are reported in Tables 4 and 5. Within the returning subset, children’s performance improved significantly from phase 1 to phase 2 (see Table 5).

**Theory of Mind Scale.** Mean scores and standard deviations of performance on this task are presented in Tables 4 and 5. In line with previous research (e.g., Wellman & Liu, 2004), the number of children who answered correctly on the 3 tasks declined with each successive task at phase 1 in the full sample. However, as can be seen in Table 5, this pattern of stepwise decline in performance was not entirely replicated in the returning subset. Within the returning subset, children’s performance on the theory of mind scale significantly improved from phase 1 to phase 2 (see Table 5).
Figure 5. Histogram and Q-Q Plot Illustrating Experience Aggregate Scores in the Full Sample at Phase 1 (n = 81)
Figure 6. Histogram and Q-Q Plot Illustrating Experience Aggregate Scores in the Returning Subset (n = 41)
Preliminary Analyses

Recall that participants who were eligible to participate in phase 2 of the study were invited to return 6 months following the initial study sessions. Phase 2 sessions ultimately took place 6 to 8 months following the phase 1 sessions, according to each family’s availability ($M = 6.39$ months, $s = .67$). Duration of time between sessions was not significantly correlated with performance on our key measures at phase 2 ($rs = -.03$ to .27, $ps = .091$ to .839).

Subsequent analyses tested differences in phase 1 task performance between children who participated in the second phase of testing (i.e., those in the returning subset), and those who did not. Recall that only children who performed poorly on the initial false belief battery were eligible to participate in the second phase of testing. In line with this criterion, children in the returning subset showed poorer performance as a group on the phase 1 false belief battery in comparison to those who participated in phase 1 only, $t (79) = 5.26$, $p < .001$, $d = 1.18$. In comparison to children who participated in phase 1 only, returning subset children also tended to be younger, $t (79) = 2.22$, $p = .029$, $d = 0.49$, score significantly lower on the PPVT-IV, $t (79) = 2.16$, $p = .034$, $d = 0.48$, and obtain lower scores on the theory of mind scale, $t (79) = 2.65$, $p = .010$, $d = 0.59$. Across groups, there were no significant differences in levels of relevant experience, $t (79) = 0.88$, $p = .382$, or RC-EF skills, $t (79) = 1.85$, $p = .068$.

Independent samples $t$-tests were conducted to examine sex differences in our measures. Although there were no sex differences in performance on our focal measures, when examining our control variables two significant differences emerged: at phase 1, girls
significantly outperformed boys on the theory-of-mind scale (full sample: \( t (79) = 3.32, p = .001 \), returning subset: \( t (39) = 2.38, p = .022, d = .75 \)), and at phase 2, boys scored significantly higher than girls on the PPVT-IV, \( t (39) = 2.37, p = .023, d = .74 \).

Another set of analyses was conducted to test the relation between age and our dependent measures. Because this study involved a narrow age range, I did not expect to find pronounced age effects in our data. Nevertheless, age was significantly correlated with false belief battery performance at phase 1 in the full sample, \( r (79) = .23, p = .044 \). Surprisingly, age was also negatively correlated with our experience aggregate at phase 2, \( r (39) = -.31, p = .046 \), indicating that younger children had higher experience aggregate scores. Follow-up analyses revealed that this effect was driven by younger children being more likely to have child-aged siblings, \( r (39) = -.33, p = .033 \), with age not significantly related to parental mental state talk, \( r (39) = -.09, p = .591 \).

**Inter-Correlations Among Measures**

Tables 6 and 7 display inter-correlations among measures within our full sample (phase 1), and in our returning subset (phases 1 and 2) respectively. Different correlation coefficients were calculated based on whether the variables included were continuous or dichotomous. That is, Pearson correlation coefficients were calculated in cases where both variables were continuous, phi coefficients were calculated in cases where both variables were dichotomous, and point-biserial correlation coefficients were calculated in cases where one variable was dichotomous and one was continuous. In our full sample at phase 1, all variables were correlated as expected based on prior research. This gives us further confidence that our sample was typical, and that our key measures were behaving as expected. In our returning subset at phase 1, false belief battery scores were significantly
correlated with RC-EF skills, but not significantly correlated with PPVT-IV or our relevant experience aggregate. In our returning subset at phase 2, false belief battery scores were significantly correlated with RC-EF and PPVT-IV, but not relevant experience.

Table 6
Inter-correlations Among Key Measures in Full Sample at Phase 1 ($n = 81$)

<table>
<thead>
<tr>
<th></th>
<th>False Belief Battery</th>
<th>RC-EF Aggregate</th>
<th>Experience Aggregate</th>
<th>PPVT Aggregate</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC-EF Aggregate</td>
<td>.57**(^a)</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experience Aggregate</td>
<td>.29**(^a)</td>
<td>.14(^c)</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>PPVT</td>
<td>.49**(^b)</td>
<td>.51**(^a)</td>
<td>.30**(^a)</td>
<td></td>
</tr>
<tr>
<td>ToM Scale</td>
<td>.32**(^b)</td>
<td>.32**(^a)</td>
<td>.06(^a)</td>
<td>.20(^b)</td>
</tr>
</tbody>
</table>

Note: * $p < .05$, ** $p < .01$

\(^a\) Point Biserial Correlation Coefficient, \(^b\) Pearson Correlation Coefficient, \(^c\) Phi Coefficient
Table 7

Inter-correlations Among Key Measures in Returning Subset at Phases 1 and 2 (n = 41)

<table>
<thead>
<tr>
<th></th>
<th>False Belief Battery</th>
<th>RC-EF Aggregate</th>
<th>Experience Aggregate</th>
<th>PPVT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Phase 1</td>
<td>Phase 2</td>
<td>Phase 1</td>
<td>Phase 2</td>
</tr>
<tr>
<td>RC-EF Aggregate</td>
<td>.36*a</td>
<td>.40*a</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Experience Aggregate</td>
<td>.10*a</td>
<td>.11*a</td>
<td>.07*c</td>
<td>-.08*c</td>
</tr>
<tr>
<td>PPVT</td>
<td>.11*b</td>
<td>.49**b</td>
<td>.52**a</td>
<td>.42**a</td>
</tr>
<tr>
<td>ToM Scale</td>
<td>-.10*b</td>
<td>.13*b</td>
<td>.24*a</td>
<td>.10*a</td>
</tr>
</tbody>
</table>

Note: * p < .05, ** p < .01

*a Point Biserial Correlation Coefficient, *b Pearson Correlation Coefficient, *c Phi Coefficient
Focal Analyses

Focal analyses were conducted with three general goals in mind. The first was to examine the concurrent interactive effect of RC-EF skills and experience on false belief understanding. The second was to examine these interactive effects longitudinally. Finally, control analyses were conducted to test alternative explanations for any observed concurrent and longitudinal interactive relations among our constructs of interest. The results of analyses addressing each of these goals are detailed in turn below.

**Concurrent relations.** To test whether there was an interactive effect of RC-EF skills and experience on false belief understanding concurrently, I examined relations between these variables within our phase 1 and phase 2 samples independently. A 2 x 2 ANOVA on data from the full sample at phase 1 revealed a main effect of RC-EF on false belief battery task performance, with the high RC-EF group out-performing the low RC-EF group, $F(1, 77) = 37.81, p < .001$, $\text{Partial } \eta^2 = .33$. There was also a significant main effect of experience, with children in the high experience group outperforming children in the low experience group on the false belief battery, $F(1, 77) = 4.48, p = .038$, $\text{Partial } \eta^2 = .06$.

Most relevant to our research questions was the interaction between RC-EF and experience. At phase 1, analyses revealed that the interactive effect of RC-EF and experience on false belief task performance was not significant, $F(1, 77) = 2.37, p = .127$, $\text{Partial } \eta^2 = .03$.

Nevertheless, given the study’s *a priori* hypothesis, I opted to run simple effects analyses to more directly test for the specific pattern of predicted associations. I first examined false belief task performance among children in the high RC-EF group separately. In line with expectations, these analyses revealed that within the high RC-EF group, children with high levels of experience ($M = 2.65$, $SE = .19$) significantly outperformed their peers ($M = 1.83$, $SE = .16$).
SE = .22) on the false belief battery, $F (1, 77) = 8.06$, $p = .006$, $Partial \eta^2 = .10$. In contrast, within the low RC-EF group there was no significant difference between the high and low experience groups on false belief battery performance, $F (1, 77) = 0.14$, $p = .709$, $Partial \eta^2 < .01$ (see Figure 7). These results suggest that experience only affected the false belief performance of children with advanced RC-EF skills.

**Figure 7.** Two-way ANOVA showing the concurrent interaction between RC-EF and relevant experience in predicting false belief (FB) understanding at phase 1. Error bars represent standard error.

I then performed a 2 x 2 ANOVA to examine concurrent relations within the phase 2 data, again with RC-EF skills and relevant experience as the independent variables and false belief battery scores as the dependent variable. At phase 2, analyses again revealed a main effect of RC-EF, with high RC-EF children performing significantly better than the low RC-EF group on the false belief battery, $F (1, 37) = 7.69$, $p = .009$, $Partial \eta^2 = .17$. 

67
However, in contrast to the phase 1 data, there was no significant main effect of experience in predicting phase 2 false belief understanding, $F(1, 37) = 0.25, p = .617$, $\text{Partial } \eta^2 = .01$. In this case, the interactive effect of phase 2 RC-EF and experience on phase 2 false belief task performance was significant, $F(1, 37) = 4.59, p = .039$, $\text{Partial } \eta^2 = .11$ (see Figure 8). Similar to phase 1, simple main effects analyses revealed that within the high RC-EF group, children with high levels of experience ($M = 3.00, SE = .27$) had significantly higher false belief understanding than the low experience children ($M = 2.19, SE = .26$), $F(1, 37) = 4.53, p = .04$, $\text{Partial } \eta^2 = .11$. Conversely, within the low RC-EF group there was no significant difference in false belief battery scores between high and low experience groups, $F(1, 37) = 1.10, p = .302$, $\text{Partial } \eta^2 = .03$. Taken together, these findings provide evidence that within this sample, the relation between relevant experience and false belief understanding only existed among children with advanced RC-EF skills.

![Figure 8](image-url)  

**Figure 8.** Two-way ANOVA showing the concurrent interaction between RC-EF and relevant experience in predicting false belief understanding at phase 2. Error bars represent standard error.
In line with the emergence account, one possibility was that the concurrent interactive relations observed were in fact a corollary of consistency in RC-EF skills over time. In other words, the observed concurrent relations among measures may not have been independently meaningful, instead reflecting longitudinal relations among variables of interest. Indeed, RC-EF skills across phases were significantly correlated, $\Phi (39) = .48, p = .002$. To test this possibility, I re-ran the phase 2 concurrent analyses as a 2 x 2 ANCOVA, controlling for phase 1 RC-EF skills. When doing so, the concurrent interaction remained significant, $F (1, 36) = 4.51, p = .041$, $Partial \eta^2 = .111$. However, while the pattern of findings was similar to our uncontrolled analyses, the high RC-EF simple effects analysis dropped below significance; within the high RC-EF group 1, the high experience children ($M = 2.88, SE = .29$) no longer significantly outperformed the low experience children ($M = 2.135, SE = .26$) on the false belief battery, although the effect was near significant, $F (1, 36) = 3.90, p = .056$, $Partial \eta^2 = .098$. In line with the uncontrolled analyses, there was no evidence for an effect of experience within the low RC-EF group, $F (1, 36) = 1.28, p = .266$, $Partial \eta^2 = .034$. In sum, these results provide preliminary evidence that concurrent interactive relations exist independent of longitudinal ones, as the component process account would suggest.

**Longitudinal relations.** I then conducted a 2 x 2 ANOVA to test whether there was an interactive effect of phase 1 RC-EF skills and experience on phase 2 false belief understanding. Results showed a significant main effect of phase 1 RC-EF on phase 2 false belief task performance, $F (1, 37) = 6.95, p = .012$, $Partial \eta^2 = .16$, with high RC-EF children significantly outperforming low RC-EF children. There was no significant main effect of experience on false belief task performance, $F (1, 37) = 0.49, p = .490$, $Partial \eta^2$
Most relevant to the study’s hypotheses, analyses revealed a significant interaction between phase 1 RC-EF and experience in predicting phase 2 false belief battery performance, $F(1, 37) = 9.48, p = .004, \text{Partial } \eta^2 = .20$ (see Figure 9). In keeping with our concurrent findings, simple main effects analyses showed that among children with high RC-EF skills at phase 1, the high experience group ($M = 3.14, SE = .28$) significantly outperformed the low experience group ($M = 2.06, SE = .31$) on the phase 2 false belief battery, $F(1, 37) = 6.92, p = .012, \text{Partial } \eta^2 = .16$. In contrast, there was no significant main effect of experience within the low RC-EF group, $F(1, 37) = 2.92, p = .096, \text{Partial } \eta^2 = .07^1$. 

---

1 Note that this main effect approached standard significance levels, but was in the opposite direction. That is, among children with low RC-EF skills, those with low levels of relevant experience ($M = 2.18, SE = .28$) performed near significantly better on the false belief battery compared to those with high levels of experience ($M = 1.50, SE = .29$). This provides further evidence that relevant experience was uniquely beneficial to children with advanced RC-EF abilities.
Figure 9. Two-way ANOVA showing the longitudinal interaction between phase 1 RC-EF and relevant experience, predicting phase 2 false belief understanding. Error bars represent standard error.

These findings suggest that phase 1 RC-EF skills influenced the extent that exposure to relevant experience impacted phase 2 false belief understanding. However, one possible alternative explanation is that the longitudinal findings described above were a corollary of consistency in RC-EF skills over time, rather than evidence of developmental relations. As previously noted, phase 1 and 2 RC-EF scores were significantly correlated, Φ (39) = .48, p = .002. Thus, phase 2 RC-EF may have been serving as a reasonable proxy for phase 1 RC-EF, leading to longitudinally significant findings that in fact reflected concurrent relations among measures. To statistically test this possibility, I re-ran the focal longitudinal analyses as a 2 x 2 ANCOVA, this time controlling for phase 2 RC-EF skills. In this case, the interaction remained significant, F (1, 36) = 9.49, p = .004, Partial η² = .21
and the pattern of simple main effects remained the same: among children with high RC-EF skills at phase 1, the high experience children ($M = 3.01, SE = .28$) significantly outperformed the low experience children ($M = 1.89, SE = .31$) on the phase 2 false belief battery, $F (1, 36) = 7.93, p = .008$, Partial $\eta^2 = .18$. In contrast, there was no significant main effect of experience within the low RC-EF group, $F (1, 36) = 2.27, p = .140$, Partial $\eta^2 = .06$. These findings suggest that RC-EF skills and experience interact to longitudinally predict false belief understanding, independent of concurrent RC-EF skills.

Following from previous research (e.g., Carlson et al., 2004; Hughes, 1998), I then tested whether the longitudinal relation reported above was in fact unidirectional in nature. A 2 x 2 ANOVA tested the reverse longitudinal interactive relation between our constructs of interest; namely, the interactive effect of phase 2 RC-EF skills and experience on phase 1 false belief understanding. As predicted, the reverse longitudinal interaction was not significant, $F (1, 37) = 1.12, p = .296$, Partial $\eta^2 = .03$. Further, there was no significant effect of experience within the high RC-EF group $F (1, 37) = 1.66, p = .206$, Partial $\eta^2 = .04$, or the low RC-EF group, $F (1, 37) = 0.11, p = .743$, Partial $\eta^2 < .01$. Taken together, these findings suggest the existence of a longitudinal, unidirectional relation between measures, whereby early RC-EF skills and experience interact to predict false belief understanding 6 months later.

**The role of age, vocabulary skills, and initial theory-of-mind knowledge.**

Although our analyses were consistent in showing that experience and RC-EF interact to affect false belief understanding, potential confounds were the extent that this interaction could be driven by relations with other related variables, such as age, precursor theory-of-mind knowledge, or vocabulary abilities. Recall that each of these factors was related to
one or more of our key measures across sessions (see Tables 3 and 4). With this in mind, I re-ran our focal analyses as 2 x 2 ANCOVAs, this time including age, theory-of-mind scale performance, and PPVT-IV scores as control variables. As can be seen in Table 9, the concurrent findings at phase 1 showed the same pattern as the initial, uncontrolled ANOVA at phase 1: the interaction between RC-EF and experience was not significant, but simple effects analyses revealed the hypothesized relations. When examining concurrent relations at phase 2 with our control variables accounted for, the interaction dropped below standard significance levels, but the pattern of results and magnitude of the effects remained similar. Thus, while it could be that the concurrent interaction at phase 2 was driven by the effects of control variables, it appears more likely that it dropped below standard significance levels here because of the reduced power associated with a fully-controlled analysis.

I then re-ran the longitudinal analysis described above as a 2 x 2 ANCOVA to examine the effects of control variables on longitudinal interactions between our variables of interest. As can be seen in Table 9, the longitudinal interactive effect of phase 1 RC-EF and experience on phase 2 false belief task performance remained strong and significant when the control variables were added to the analysis. In line with the uncontrolled analyses, simple effects analyses revealed that experience was only significantly related to false belief task performance among children with advanced RC-EF skills.
### Table 8. Results of Two-Way ANCOVAs Controlling for Age, Theory-of-Mind Scale scores, and PPVT scores

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Control Variables</th>
<th>RC-EF x Experience interaction effect</th>
<th>Simple Effects Analyses (Experience within EF group)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Concurrent Analyses Predicting False Belief Performance</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Phase 1 RC-EF         | Phase 1 PPVT      | $F (1, 74) = 1.58, p = .212, \textit{Partial } \eta^2 = .02$ | High RC-EF:  
  $F (1, 74) = 4.86, p = .031, \textit{Partial } \eta^2 = .06$  
  Low RC-EF:  
  $F (1, 74) = .081, p = .777, \textit{Partial } \eta^2 < .01$ |
| Phase 1 Experience    | Phase 1 ToM Scale | $F (1, 74) = 4.86, p = .031, \textit{Partial } \eta^2 = .06$ | High RC-EF:  
  $F (1, 74) = .081, p = .777, \textit{Partial } \eta^2 < .01$ |
| Phase 1 PPVT          | Phase 1 ToM Scale | $F (1, 74) = 4.86, p = .031, \textit{Partial } \eta^2 = .06$ | High RC-EF:  
  $F (1, 74) = .081, p = .777, \textit{Partial } \eta^2 < .01$ |
| Phase 1 age           | | | |
| Phase 2 RC-EF         | Phase 2 PPVT      | $F (1, 34) = 2.34, p = .136, \textit{Partial } \eta^2 = .06$ | High RC-EF:  
  $F (1, 34) = 3.70, p = .063, \textit{Partial } \eta^2 = .10$  
  Low RC-EF:  
  $F (1, 34) = 0.10, p = .756, \textit{Partial } \eta^2 < .01$ |
| Phase 2 Experience Aggregate | Phase 2 ToM Scale | $F (1, 34) = 2.34, p = .136, \textit{Partial } \eta^2 = .06$ | |
| Phase 2 PPVT          | Phase 2 ToM Scale | $F (1, 34) = 2.34, p = .136, \textit{Partial } \eta^2 = .06$ | |
| Phase 2 age           | | | |
| **LongitudinalAnalyses – Phase 1 IVs Predicting Phase 2 False Belief** | | | |
| Phase 1 RC-EF         | Phase 1 PPVT      | $F (1, 34) = 10.45, p = .003, \textit{Partial } \eta^2 = .24$ | High RC-EF:  
  $F (1, 37) = 7.90, p = .008, \textit{Partial } \eta^2 = .19$  
  Low RC-EF:  
  $F (1, 37) = 2.47, p = .125, \textit{Partial } \eta^2 = .07$ |
| Phase 1 Experience Aggregate | Phase 1 ToM Scale | $F (1, 34) = 10.45, p = .003, \textit{Partial } \eta^2 = .24$ | |
| Phase 1 PPVT          | Phase 1 ToM Scale | $F (1, 34) = 10.45, p = .003, \textit{Partial } \eta^2 = .24$ | |
| Phase 1 age           | | | |

Note. ToM = Theory of Mind, PPVT = Peabody Picture Vocabulary Test
These findings suggest that, at the very least, the observed longitudinal relations between our variables of interest are not driven by the effects of age, vocabulary skills, or initial theory-of-mind knowledge. Nevertheless, given that neither concurrent interaction was statistically significant when control variables were included, I opted to conduct exploratory analyses to further examine the role played by each control variable. These analyses aimed to address what I deemed to be the three most plausible alternative explanations for our concurrent findings involving our control variables. The first was that RC-EF skills were not themselves important for concurrent false belief task performance, but instead were acting as a rough proxy for vocabulary skills. The second was that it was not RC-EF skills but precursor theory-of-mind knowledge that was critical in predicting false beliefs task performance. The third was that RC-EF skills were acting as a rough proxy for general, age-related maturation, which itself was the more important factor for false belief task performance. To test these alternative explanations for our concurrent findings, I re-ran the concurrent ANOVAs, each time replacing RC-EF skills with one of our control variables (PPVT performance, theory-of-mind scale scores, and age, in turn). None of these analyses revealed a significant interactive effect of the control variables and experience on false belief understanding (Phase 1: $F_s = 0.03$ to $0.17$, $p_s = .678$ to .853, Phase 2: $F_s = 0.08$ to $3.07$, $p_s = .088$ to .773). Thus, although the concurrent interaction effect of RC-EF and experience on false belief understanding dropped below significance in our fully controlled analyses, I do not have evidence that any of our control variables
more effectively determine the extent that children benefit from experience in the service of false belief task performance.
Chapter 4

Discussion

A growing body of work has identified a relation between RC-EF skills and false belief reasoning in young children (e.g., Carlson & Moses, 2001), yet there is still debate regarding the nature of this relation. This longitudinal study aimed to test the emergence account’s hypothesis that RC-EF skills facilitate learning from relevant experience to develop a conceptual understanding of false beliefs over the preschool years. This study’s findings are summarized below, with concurrent and longitudinal analyses covered in turn.

Concurrent Relations

Both the phase 1 and 2 data showed initial patterns of results that were consistent with the hypothesized interactive relation between RC-EF and relevant experience in influencing false belief performance. Although the interaction was only statistically significant in the phase 2 sample, simple effects analyses at both time points revealed that children with advanced RC-EF skills showed a positive effect of relevant experience on false belief understanding. As expected, this effect was not paralleled in the low RC-EF sample.

Further analyses examined the possible effect of consistency in RC-EF skills over time on concurrent interactive relationships. Results showed that at phase 2, RC-EF and experience interacted to concurrently affect false belief understanding, even after controlling for phase 1 RC-EF skills. However, follow-up analyses showed that within the high RC-EF group, the effect of experience on false belief understanding dropped to near
significant levels. Thus, while a concurrent interactive relation between variables appears to remain independent of longitudinal RC-EF effects, the simple effects were weaker than in uncontrolled analyses.

A second set of control analyses tested whether relations with age, vocabulary abilities, or initial theory-of-mind knowledge could explain concurrent interactions among measures. While the general pattern of concurrent findings remained consistent when these variables were controlled, neither the phase 1 nor phase 2 interactions were statistically significant. Follow-up analyses showed that, in line with the uncontrolled analyses, at phase 1 the high RC-EF children showed an effect of experience on false belief understanding that was not paralleled in the low RC-EF group. However, at phase 2 this effect dropped slightly below standard significance levels.

Given that controlling for age, vocabulary abilities and initial theory-of-mind impacted the strength of concurrent interactions, exploratory analyses were conducted to test whether the control variables themselves interacted with experience to predict false belief task performance. None of these analyses revealed a significant interaction, failing to provide evidence for the possibility that RC-EF skills were acting as a proxy for these variables in our initial concurrent analyses. Taken together, these findings offer preliminary evidence that RC-EF skills and relevant experience interact to concurrently predict false belief understanding when relevant control variables are taken into account.

The emergence account suggests that RC-EF skills are necessary for children to effectively capitalize on experiences relevant to developing an understanding of false belief
concepts. From this account, concurrent interactive relations between RC-EF and experience would be hypothesized only as a byproduct of the prolonged, longitudinal effects of these variables on subsequent false belief understanding. In line with the emergence account, concurrent interactive effects were weakened when relevant variables were controlled. Nevertheless, I did gain some evidence for a phase 2 concurrent interaction that was independent of the longitudinal effects of RC-EF. These findings are not easily explained by the emergence account.

Instead, this pattern of concurrent results is more straightforwardly in line with the component process account. The component process account suggests that RC-EF skills play an important role in online false belief reasoning by facilitating self-perspective inhibition. According to this account, children (and adults) must have sufficient RC-EF skills to successfully inhibit their own perspective in order to effectively reason about the contents of other minds. This account leaves open the possibility that false belief reasoning is also influenced by factors independent of RC-EF skills, suggesting that RC-EF skills are necessary but not sufficient for false belief reasoning.

Let us suppose, then, that RC-EF skills are important for online false belief reasoning via self-perspective inhibition. Let us further suppose that relevant experience is independently important for developing knowledge of other minds, which is necessary for other perspective taking. From this account, then, even if children have sufficient RC-EF skills for self-perspective inhibition, they may not have developed the underlying conceptual knowledge necessary for other perspective taking if they have not been privy to
high levels of relevant experience. If this framework accurately captures the roles of RC-EF and experience in false belief reasoning, I would therefore expect children with both advanced RC-EF skills and high levels of false-belief-relevant experience to outperform their peers on concurrently measured false belief tasks. Although concurrent interactions were weakened when longitudinal consistency in RC-EF was controlled, there was some evidence for this pattern of results. This provides evidence in favour of the component process account of the relation between RC-EF and false belief reasoning.

**Longitudinal Relations**

With respect to longitudinal findings, analyses revealed that early RC-EF skills interacted with false-belief-relevant experience to predict false belief understanding 6 months later. Interestingly, this longitudinal effect explained twice as much variance as the concurrent phase 1 effect noted above (as measured by partial eta squared), and 6 times more than the concurrent effect at phase 2. As expected, follow-up analyses revealed that only children with high RC-EF skills showed a significant, positive effect of false-belief-relevant experience on later false belief understanding. Further analyses suggest that the longitudinal, interactive relation is unidirectional in nature, as phase 2 RC-EF skills and relevant experience did not interact to predict phase 1 false belief understanding. This gives some confidence that the longitudinal, predictive interaction between RC-EF and experience was driven by the role that these variables play in promoting false belief development.
Control analyses involving longitudinal interactive relations were also clearly in line with the study’s hypotheses. First, there was an interactive effect of early RC-EF skills and experience on later false belief understanding, even when controlling for later RC-EF skills. Thus, the longitudinal interactive pattern of findings existed independent of concurrent relations between RC-EF and false belief understanding. Second, longitudinal interactive relations remained significant when age, vocabulary skills, and initial theory-of-mind knowledge were controlled for. Thus, relations with these factors cannot account for the focal longitudinal findings. Taken together, these results provide strong, consistent support for the hypothesis that RC-EF skills and relevant experience interact to longitudinally predict false belief understanding in preschool-aged children.

Overall, this study’s longitudinal results provide evidence for the emergence account of the relation between false belief understanding and RC-EF skills. The longitudinal interaction that emerged in these data was straightforwardly predicted by this framework; only children with advanced RC-EF skills were expected to show benefits from false-belief-relevant experience, as RC-EF skills are hypothesized to be necessary for experience-driven social conceptual change. Moreover, this longitudinal interactive effect was expected to emerge independent of relevant control variables, including the effects of later RC-EF skills on false belief performance.

Conversely, these longitudinal findings are not easily accounted for by the component process account. If RC-EF skills exert their effect through self-perspective
inhibition when children are actively engaging in false belief reasoning, then there would be no clear reason to expect this developmental interaction with experience to emerge.

**Can a Case be Made for the Early Understanding Account?**

Another relevant question concerns how these findings might relate to the early understanding account of relations between RC-EF and false belief performance. Recall that the early understanding account proposes RC-EF to be superficially related to false belief task performance, with children requiring a certain level of RC-EF skills to demonstrate underlying mental state knowledge on standard false belief tasks (e.g., Leslie, 2005). Proponents of this account also tend to suggest that generating the content of others’ false beliefs is an automatic process, with RC-EF skills involved later, at the point when children must actively select a less salient response or perspective in false belief reasoning contexts. From this viewpoint, there is no straightforward reason to expect an interaction between RC-EF skills and experience in predicting false belief understanding; in fact, there is no reason to expect any role for relevant experience in preschool false belief performance shifts. There is also no reason to expect relations with RC-EF skills to exist longitudinally, over and above concurrent relations. Thus, while prior research has made it clear that standard false belief reasoning contexts do indeed pose non-trivial RC-EF demands, it does not appear that these demands can account for the present study’s pattern of longitudinal, interactive relations.
An Integrated Theoretical Approach to Understanding the Relation between RC-EF Skills and False Belief Reasoning

As discussed above, finding concurrent interactive relations among measures—even when controlling for longitudinal consistency in RC-EF—is process account, but not the emergence account. Conversely, finding longitudinal interactive effects that are independent of relevant controls aligns well with the emergence account, but not the component process account. It is possible, then, that both the emergence account and the component process account capture different aspects of the relation between RC-EF skills and false belief understanding. While these accounts have been discussed separately to this point, they are not inherently mutually exclusive; some aspects of the two theories may in fact dovetail nicely. For instance, disengaging from the self-perspective could be critical for both effectively engaging in online false belief reasoning (in line with the component process account), and also for learning about other minds from experience (in line with the emergence account). Figure 10 illustrates an integrated framework that incorporates elements from both component process and emergence accounts to characterize the relation between RC-EF skills and false belief reasoning. In the following section I more thoroughly address possible roles for RC-EF skills in developing a conceptual understanding of other minds.
Figure 10. An integrated model of the mechanisms underlying the relation between RC-EF skills and false belief reasoning in preschool-aged children

Mechanisms Underlying the Relation between RC-EF and Experience-Driven Social Conceptual Development

According to the theory-theory, preschoolers’ conceptual understandings of other minds develop from relevant experiences, namely interactions with others. The emergence account adds to this theory by suggesting an important role for RC-EF skills in the process of developing conceptual knowledge from experience. An important question that emerges from this theory concerns the specific mechanisms through which RC-EF skills exert an effect on social-conceptual change over time. Several authors have suggested that RC-EF skills may equip children with the tools necessary to more regularly engage in social interactions, which in turn provide a rich source of information from which to glean
information about other minds (e.g., Hughes, 1998). Indeed, several studies have implicated executive functioning skills in social competence, suggesting that social interactions rely upon the successful regulation of behaviour (e.g., Fahie & Symons, 2003; Hughes, Dunn, & White, 1998; Razza & Blair, 2009). Thus, children with advanced RC-EF skills may ultimately benefit from more frequent and successful social interactions, which in turn provide more opportunities to learn about other minds.

Another way in which RC-EF skills may contribute to social conceptual development is by facilitating theory change from relevant experience. There are at least three mechanisms through which RC-EF skills may exert such an effect. First, RC-EF skills may enable children to consider and attend to false-belief-relevant information (Benson et al., 2013). Indeed, two children in the same situation may take very different information from their experience, and whether they absorb false-belief-relevant information may be dependant on their RC-EF skills. One reason this may be true is if RC-EF skills help children to ignore irrelevant distractions in their environment (Benson et al., 2013). However, even when children are not outwardly distracted, they may have difficulty attending to others’ false beliefs because doing so requires self-perspective inhibition (e.g., Carlson & Moses, 2001; Samson et al., 2005). Similarly, recognizing another person’s false belief may necessitate separating or “distancing” mental representations from objective reality (Moses & Sabbagh, 2007). In sum, low RC-EF children may face significant challenges in effectively attending to and absorbing false-belief-relevant information.
A second way that RC-EF skills may facilitate learning about other minds from experience is by enabling children to explicitly reflect upon discrepancies between prior expectations and subsequent outcomes (Benson et al., 2013; Zelazo, 2004). Children with low RC-EF skills may not be able to disengage from the present to actively reflect upon instances where their expectations regarding others’ actions are in conflict with what ultimately transpires.

Finally, RC-EF skills may facilitate the process of updating previously-established false-belief-relevant theories (Benson et al., 2013). If children have an initial bias to assume that other individuals share their own personal perspectives and beliefs, then advances in understanding other minds would necessarily involve updating this prior theory. Indeed, research suggests that updating theories based on new, incongruent evidence appears to be particularly difficult for young children (e.g., Siegler & Chan, 1998). Thus, preschool-aged children may be reliant on RC-EF skills when challenged to integrate new, false-belief-relevant feedback with previously-established beliefs about other minds.

**Study Implications**

This study’s findings provide new insights into the ways in which factors that are important for social-cognitive functioning exert their effect. Most notably, these findings constitute the first direct evidence for the emergence account’s theory that RC-EF skills interact with experience over time to predict subsequent advances in false belief reasoning. In line with the component process account, results further indicate a facilitative role for RC-EF skills in the process of online false belief reasoning. Taken together, the observed
pattern of concurrent and longitudinal results suggests a complex relation between RC-EF and false belief reasoning, inspiring an integrated theoretical approach to characterizing the role that RC-EF skills and relevant experience play in social cognition.

Beyond theoretical implications, the results of this study have noteworthy real-world implications for parents and educators. This study’s findings suggest that RC-EF skills are necessary for both the development and effective application of social-conceptual knowledge. With this in mind, those who wish to promote social abilities in young children may choose to encourage and support RC-EF skills to this end. Research has shown that children’s executive functioning skills can be improved through classroom interventions designed to give children regular practice exerting their self-control (e.g., Diamond, Barnet, Thomas, & Munro, 2007). These findings, in conjunction with the present study’s results, suggest that RC-EF training programs may offer a means of positively impacting social functioning if used in conjunction with more conventional interventions that provide socially-relevant information.

Results of this study also offer insights into the social functioning of clinical populations with executive functioning deficits. For instance, Attention-Deficit/Hyperactivity Disorder (ADHD) is a developmental disorder characterized by inattention and/or hyperactivity-impulsivity (American Psychiatric Association, 2013). While the disorder centers on executive functioning deficits, this study suggests that children with ADHD may also be prone to difficulties in social domains. Specifically,
these children may not be equipped to advance their mental state knowledge as efficiently as typically-developing children, and they may also have difficulty making use of that knowledge in their day-to-day lives. In line with this prediction, a growing body of research is revealing that children with ADHD do indeed demonstrate deficits in social functioning (e.g., Fahie & Symons, 2003, see Benson & Sabbagh, in press, for a review). With a better understanding of the precise mechanisms through which RC-EF skills impact social functioning, researchers and practitioners can work to develop strategies that help children with conditions like ADHD function more effectively in their daily lives.

In a similar vein, this study has implications for understanding social-cognitive deficits among low socio-economic status (SES) populations. Both RC-EF skills and false belief understanding have been shown to vary according to socio-economic status (SES), with low SES groups at a significant disadvantage when compared to middle-class populations (e.g., Bernier, Carlson, Dechenes, & Matte-Gagne, 2012; Cutting & Dunn, 1999). This study’s findings suggest that limited RC-EF skills in low-SES groups may have a deleterious effect on advances in social-cognitive knowledge. Future research is necessary to explore whether interventions that incorporate RC-EF skills training are effective in improving the social knowledge and functioning of children from low SES backgrounds.

The results of this study further highlight the importance of considering the effect of situational factors that impede executive functioning skills on children’s theory-of-mind reasoning. A number of factors can influence a child’s ability to engage his or her
executive functioning skills, including being hungry (Gailliot, 2008), physically tired (Nilsson et al., 2005), cognitively tired (Lillard & Peterson, 2001), or in a negative, emotionally-aroused state (Padmala, Bauer, & Pessoa, 2011). With this in mind, a wise strategy for those interacting with young children would be to avoid staging theory-of-mind interventions that involve presenting socially-relevant information when a child’s executive functioning skills are highly taxed. Instead, a better strategy would be to wait for a time when the child is better equipped to effectively make use of explicit instruction and guidance on mental state reasoning; doing so will improve the likelihood that this type of feedback will ultimately promote long-term social-cognitive development.

Limitations and Future Directions

There are several limitations of this study. While this study’s sample size was in line with other longitudinal projects with similar research goals (e.g., Hughes, 1998), it may not have afforded enough power to detect relatively small interaction effects (for instance, concurrent interactive relations among measures). As is the case with other studies of this nature, replication is therefore critical.

A second limitation concerns the fact that this study did not involve experimental manipulations, thus limiting the extent that causal relations can be conclusively tested. Given our interest in capturing the effects of naturalistic forms of false-belief-relevant experience on natural trajectories of false belief understanding, a study design in which variables were experimentally manipulated was not feasible. While the pattern of findings
that ultimately emerged was in line with our theoretically driven, causal hypotheses, it is worth noting that causal links between our variables of interest cannot be conclusively proven from these data. For instance, while finding links between parental mental state talk and children’s false belief reasoning suggests an effect of relevant experience on false belief development, it remains possible (although seemingly improbable) that the two factors are not causally related, instead representing a byproduct of some unaccounted-for third variables (e.g., genetic predisposition for social behaviour). While I think that alternative explanations for these findings are unlikely, it is worth noting that they remain a possibility.

A third limitation concerns this study’s longitudinal control analysis examining reverse developmental relations in the returning subset (i.e., phase 2 RC-EF and experience predicting phase 1 false belief performance). Theoretically, I did not expect this reverse longitudinal interaction to be significant, and results were in line with this prediction. However, a possible confound in this analysis was limited variability in the returning subset’s phase 1 false belief scores. Recall that phase 1 false belief battery performance was used as inclusion criteria for phase 2 of the study, with only children who obtained less than 50% on the measure qualifying to participate. As a result, the returning subset’s phase 1 false belief battery scores were restricted in variability. This measure continued to correlate with phase 1 RC-EF skills, giving some confidence that meaningful variability remained despite the restricted range of scores. Nevertheless, restricted variability may have contributed to null findings in this longitudinal control analysis.
While results from this study suggest that RC-EF skills are important for social conceptual change, this study does not provide insight into exactly how RC-EF skills exert this effect. Recently, theory-of-mind research has begun to focus on understanding exactly how fine-grained shifts in conceptual knowledge take place (e.g., Gopnik & Wellman, 2012). Future research is necessary to further test the ways in which RC-EF skills contribute to the process of learning about other minds over time.

It is also worth noting that this study examined the effects of RC-EF skills on a relatively circumscribed aspect of social-cognitive development that emerges during a narrow window of time. There are reasons to suspect that RC-EF skills play an especially important role in false belief reasoning, given its apparent reliance on self-perspective inhibition. Nevertheless, research suggests that RC-EF skills impact learning in a number of cognitive domains, including mathematics and language (e.g., Espy et al., 2004; McClelland et al., 2007). With this in mind, an important question concerns whether RC-EF skills play a role in other aspects of social-cognitive knowledge acquisition during other stages of development. Testing the specificity (or lack thereof) in the aspects of conceptual development that rely upon RC-EF may ultimately help to shed light on the precise mechanisms through which RC-EF skills exert their effect. If, for instance, RC-EF skills are primarily recruited for learning when children have a previously-established, incorrect theory to overturn, then RC-EF skills may have particular involvement in overcoming the tendency to focus on old theories in light of new evidence. In sum, further research should
be directed at better understanding the breadth of impact that RC-EF skills have on conceptual development across varying domains and stages of development.

Finally, research is necessary to examine the neuro-maturational processes related to false belief reasoning and development, with specific focus on how these processes might themselves relate to both RC-EF skills and relevant experience. Research has shown that preschoolers’ theory-of-mind reasoning is correlated with neuro-maturation in specific brain regions—in particular, the medial prefrontal cortex and temporal parietal junction (e.g., Sabbagh et al., 2009). No research to date has been directed at understanding how experiences found to be important for false belief reasoning relate to the neuro-correlates of false belief understanding, or whether RC-EF skills interact with this relation. While the present study represents an important step toward understanding the roles of both RC-EF skills and experience in false belief development, future research should aim to understand the extent that these factors are related on a neurological level.

**Conclusion**

In sum, this study represents an important addition to existing research on the relation between RC-EF skills and false belief understanding. Through the use of a longitudinal study design, I was able to examine how children’s RC-EF skills interact with exposure to naturalistic false-belief-relevant experience to facilitate improvements in performance on false belief tasks. These study results provide strong evidence that early RC-EF skills interact with experience to predict false belief understanding 6 months later. More specifically, only children with advanced RC-EF skills showed a positive effect of
high levels of experience on subsequent false belief performance. The interaction between 
RC-EF and experience remained statistically significant when potentially confounding 
variables were statistically controlled. These findings are in line with the emergence 
account, which suggests that children with well-developed RC-EF skills are better suited to 
make use of the types of experiences that are important to advance false belief knowledge 
over time.

This study also provides evidence that RC-EF skills and experience interact to 
concurrently predict false belief understanding, although the hypothesized pattern of 
findings was not consistently significant when relevant variables were controlled. The 
presence of a concurrent interactive relation that exists independent of longitudinal 
relations among measures aligns with the component process account of the relation 
between RC-EF and false belief. According to this theory, RC-EF skills are necessary for 
both children and adults to engage in false belief reasoning in all cases that involve self-
perspective inhibition. Although future work is necessary to further explore concurrent 
relations among measures, these findings suggest that an integrated approach that captures 
elements of both the emergence and component process accounts may be the most 
appropriate way to characterize the relation between RC-EF skills and false belief 
understanding. Overall, this study provides an important contribution to our understanding 
of the interactive relation between factors that can facilitate, or conversely impede, healthy 
social development.
References


Appendix A

Sample False Belief Battery Tasks

Contents Change

Show child a Smarties box with pencil crayons inside. “Here’s a Smarties box. What do you think is inside the Smarties box?” Next, the Smarties box is opened. “Well let’s see….it’s really pencil crayons inside.” Close Smarties box. “Okay, what’s in the Smarties box?” If the child responds correctly that it contains pencils, then proceed to the following paragraph. If the child responds incorrectly, start the task from the beginning.

Show child a stuffed animal, Tigger. “Here’s Tigger. Tigger’s never seen inside this Smarties box before.” Memory question: “Has Tigger seen inside this box?” If child answers incorrectly, repeat Tigger’s introduction and ask memory question again until child answers correctly. Test question: “So, what does Tigger think is in the box, Smarties or pencil crayons?” To be correct, the child must answer “Smarties” to the test question.

Location Change False Belief

Place a miniature toy box, a miniature set of drawers, and a toy car on the table in front of the child. “Now let me show you what else I have. Here are two kids. This one’s name is Eric,” (make Eric say hello) “and this is Cristyne” (make Cristyne say hello). “Eric and Cristyne are playing with this toy car.” Show the dolls playing together. “After a while, Eric decides to go play outside. He puts the car away in the toy box, and goes outside to play.” Show Eric placing the car in the toy box, and then remove him from the scene.
“While he’s gone, Cristyne wants to play with the car some more, so she gets the car out of the toy box to play.” Show Cristyne taking the car out and playing with it. “Now it’s time for Cristyne to go eat lunch. Before she goes, she puts the car away, but she puts it in this drawer. Then she goes away to eat.” Show Cristyne placing the car in the drawer, and then remove her from the scene. “Now here comes Eric again. He wants to play with the car some more.” Memory question: “So, [child’s name], did Eric see the toy car get moved?”

If the child responds correctly with “no”, say: “That’s right! He didn’t see it get moved!”, and proceed to the next paragraph. If the child responds incorrectly with “yes”, say: “Actually, he didn’t see it get moved. Let’s watch the story again.” Continue on to repeat the task from the beginning, and ask the memory question again until the child answers correctly.

Test question: “Alright, so when Eric comes back, where will he look for the car?”

To be correct, the child must answer that Eric will look in the toy box for his car.

**Appearance-Reality**

**Candle/Orange.** Place an object that looks like an orange, but is actually a candle on the table in front of the child. “Look what I have here. What does this look like?” Wait for the child to respond with ‘orange’. “Yeah, it looks like an orange, doesn't it? But really and truly it's a candle, see?” Encourage the child to touch the object. Test question 1: “When you look at this right now, does it look like a candle or an orange?” Wait for the child to respond. Test question 2: “And what is this really and truly, a candle or an
orange?” To be correct, the child must respond with “orange” to the first question and “candle” to the second question.

**Red/Black Castle.** Place a picture of a red castle in front of the child. “Now look at this. This is a red castle. But look, when I put this cover over it, it looks black, see?” Place a transparent, coloured cover over the castle to make it appear black. “So, really and truly it's red.” Uncover the castle briefly. “But we can make it look black.” Replace the cover. Test question 1: “When you look at the castle right now, does it look red or black?” Test question 2: “And what colour is the castle really and truly, red or black?” To be correct, the child must respond with ‘black’ to the first question and ‘red’ to the second question.

**Deceptive Pointing**

**Practice Trial.** Place a toy horse and two boxes on the table in front of the child. Bring out a puppet, Samantha. “Now we're gonna play a game with Samantha. See this horse? And see these boxes over here? I'm gonna put the horse inside one of the boxes like this.” Place the horse in one of the boxes. “I can point to the box with my finger like this so we'll know which box the horse is in.” Point to the box that contains the horse. “Now you try. Take the horse out of the box.” Wait for the child to remove the horse. “Now go ahead and put the horse in the other box.” Wait for the child to place the horse in the other box. “Ok, now, point to that box so we'll know where the car is.” Wait for the child to point to the box that contains the horse. “Good job! See, now we can tell where the horse is by pointing to it!”
**Test Trial 1.** “Now let's put the horse in a box and see if Samantha can find it. But first, Samantha is going to leave and go in here (toy bin to the side of the experimenter) so she can't see.” Place Samantha in the toy bin and close the lid. “Okay, go ahead and put the horse in one of the boxes.” Wait for the child to place the horse in one of the boxes. “Hey, I have a great idea! Let's play a funny trick on Samantha. Let's play a trick so she can't find the horse. Maybe we could trick her so she'll look in the wrong box, okay? Now, remember, we're gonna play a funny trick on Samantha. We can play a trick by pointing so she won't find the horse. Are you ready?” Wait until the child says he/she is ready. “Here comes Samantha!” Remove Samantha from the bin, bring her up to the boxes, and make Samantha ask: “Where’s the horse?”

If the child hesitates at first, say: “Samantha is asking where the horse is. Where do you want to point?” If the child still does not respond, say: “Which box do you want to point to?” If the second prompt is also ineffective in inducing the child to respond, say: “Do you want to point to this box (point) or this box (point)?” Wait until the child responds by pointing to one of the boxes. “Ok, Samantha. You can look now!” Make Samantha open the box that the child pointed to. If the child used deception and thus pointed to the empty box, say: “Oh, she didn’t find it! I guess we tricked her! We’re so tricky! Go ahead and show her where it really is!” If the child failed to use deception and thus pointed to the box that contained the horse, say: “Oh, she found it! I guess we didn’t trick her this time.” Place Samantha back in the bin. To be correct, the child must point to the empty box, thus conveying a misleading clue about the whereabouts of the toy.
**Test Trial 2.** Bring out a new puppet, Puppy. “Now it’s time for Puppy to come play with us. Let's put the horse in a box and see if Puppy can find it. But first, Puppy is going to leave and go in here (toy bin to the side of the experimenter) so he can't see.” Place Puppy in the toy bin and close the lid. “Okay, go ahead and put the horse in one of the boxes.” Wait for the child to place the horse in one of the boxes. “Hey, I have a great idea! Let's play a funny trick on Puppy. Let's play a trick so he can't find the horse. Maybe we could trick him so he’ll look in the wrong box, okay? Now, remember, we're gonna play a funny trick on Puppy. We can play a trick by pointing so he won't find the horse. Are you ready?” Wait until the child says he/she is ready. “Here comes Puppy!” Remove Puppy from the bin, bring him up to the boxes, and make Puppy ask: “Where’s the horse?”

If the child hesitates at first, say: “Puppy is asking where the horse is. Where do you want to point?” If the child still does not respond, say: “Which box do you want to point to?” If the second prompt is also ineffective in inducing the child to respond, say: “Do you want to point to this box (point) or this box (point)?” Wait until the child responds by pointing to one of the boxes. “Ok, Puppy. You can look now!” Make Puppy open the box that the child pointed to. If the child used deception and thus pointed to the empty box, say: “Oh, he didn’t find it! I guess we tricked him! We’re so tricky! Go ahead and show him where it really is!” If the child failed to use deception and thus pointed to the box that contained the horse, say: “Oh, he found it! I guess we didn’t trick him this time.” Place Puppy back in the bin. To be correct, the child must point to the empty box, thus conveying a misleading clue about the whereabouts of the toy.
“That was fun! You did a great job! Now remember, it’s fun to play tricks sometimes, but we should always tell the truth, okay?”
Appendix B
Executive Functioning Task Scripts

Bear-Dragon Task

“Ok, I’m going to ask you to do some silly things before we start our next game.” Proceed with the following list of commands, modeling each action along with the child. “Stick out your tongue. Touch your ears. Touch your teeth. Touch your eyes. Clap your hands. Touch your feet. Touch your head. Touch your tummy. Touch your nose. Wave your hand. Okay, good job!”

“Now I have a game that we can play with these puppets.” Bring out two puppets: a friendly-looking bear, and a mean-looking dragon. Place one on each hand, and focus the child’s attention on the bear puppet. “This puppet is a nice bear. When he talks to us, we’ll do what he tells us to do.” Focus the child’s attention on the dragon puppet. “This puppet isn’t very nice. This puppet is a dragon. When he talks to us, we won’t listen to him. If he tells us to do something, we won’t do it. Okay, let’s practice one time.” Bring the bear towards the child. “This is the good bear. He says, ‘Touch your nose’.” Use a mellow, nice voice to say the bear’s instructions. If the child does not touch his/her nose, say: “Remember, we listen to the nice bear and do what he tells us to do because that’s how we play the game.” Continue to repeat the instructions and model the action if necessary until the child succeeds. Wait until the child touches his/her nose. “That’s right! Now let’s practice with the naughty dragon. In this game, we won’t do what the dragon tells us to do
because he’s not so nice. ‘Touch your tummy’. Use a low, gruff voice to say the dragon’s instructions. If the child toches his/her stomach, say: “Remember, we’re not going to listen to the mean dragon. We won’t do what he tells us to do because that’s how we play the game.” Repeat the instructions until the child successfully refrains from touching his/her stomach. If necessary, hold the child’s hands down on the sixth try. Wait until the child succeeds. “Yeah! Good job! That was fun! So, when the bear tells you to do something, do you do it?” Wait for the child to respond and correct if necessary. “And when the dragon tells you to do something, do you do it?” Wait for the child to respond and correct if necessary. “Okay, let’s play some more!”

On the following test trials, alternate between using the dragon and bear voices. Whenever the dragon provides the instruction, bring the dragon slightly towards the child. Likewise, whenever the bear provides the instruction, bring the bear slightly towards the child. Bear: “Stick out your tongue.” Dragon: “Touch your ears.” Bear: “Touch your teeth.” Dragon: “Touch your eyes.” Bear: “Clap your hands.” Provide the following reminder, regardless of performance: “Remember the way we play this game. We do what the bear tells us to do because he’s nice, but we won’t do what the dragon tells us to do because he’s not so nice.” Dragon: “Touch your feet.” Bear: “Touch your head.” Dragon: “Touch your tummy.” Bear: “Touch your nose.” Dragon: “Wave your hand."

On each of the 5 bear test trials, code children’s behavioural responses as follows: 0 = failure to move, 1 = a wrong movement, 2 = a partial correct movement, 3 = a full correct movement. On each of the 5 dragon test trials, code children’s behavioural responses as
follows: 0 = full correct movement, 1 = partial correct movement, 2 = incorrect movement (e.g., touches feet when instructed to touch head), 3 = strategic movement (e.g., shakes head no, moves hands behind back), and 4 = no movement.

**Grass/Snow Task**

On the table in front of the child, place a board that has a solid white card attached to the upper left corner, a solid green card attached to the upper right corner (both cards are 15 x 10 cm), and two fabric cut-outs shaped like a child’s hands centered below the cards. Place the child’s hands on top of the felt child-sized hands on the board. “Now we’re going to play a game with this board. Do you know what colour grass is?” Wait for the child to respond with ‘green’, and correct if necessary. “That’s right! And do you know what colour snow is?” Wait for the child to respond with ‘white’, and correct if necessary. “Very good. We’re going to play a silly game. In this game, when I say the word ‘grass’, I want you to point with your finger to the white card, like this.” Demonstrate pointing to the white card. “Can you point to the white card?” Praise the child if he/she pointed correctly, prompt if not. “And when I say the word ‘snow’, I want you to point with your finger to the green card like this.” Demonstrate pointing to the green card. “Can you point to the green card?” Praise the child if he/she points correctly, prompt if not.

Begin the practice trials. “Grass.” Wait for the child to respond. If the child hesitates, say: “What card do you want to point to for this one?” Praise the child if he/she responds correctly. If he/she is incorrect, say: “Remember, this is a silly game. When I say ‘grass’, I want you to point over here to the white card. When I say snow, that’s when you
point to the green card. Grass.” Continue reiterating the instructions until the child responds correctly. “Very good! Snow.” Wait for the child to respond. If the child hesitates, say: “What card do you want to point to for this one?” Praise the child if he/she responds correctly. If he/she responds incorrectly, say: “Remember, this is a silly game. When I say ‘snow’, I want you to point over here to the green card. When I say grass, that’s when you point to the white card. Snow.” Continue with the practice trials until the child correctly responds to both trials consecutively.

Begin the test trials. If the child hesitates, say: “What card do you point to for this one?” Do not give feedback on the test trials. The order of the 16 test trials is as follows: Grass, Snow, Snow, Grass, Snow, Grass, Snow, Grass, Grass, Snow, Grass, Grass, Grass, Snow, Grass, Snow. On each trial, the child will receive a score of 1 if he/she initially points to the correct card, and a score of 0 if he/she initially points to the incorrect card, or to both cards at once.

**Card Sort Task (Dimensional Change Card Sort)**

Place 2 boxes on the table in front of the child. The box to the left has a red rabbit card pasted on it, while the box to the right has a blue boat card pasted on it. “Now we’re going to play a game. This is the shape game. All the rabbits go in this box.” Point to the left red rabbit box. “And all the boats go in that box.” Point to the right blue boat box. “We don’t put any rabbits in that box. No way! We put all the rabbits over here.” Point to the left red rabbit box. “And only boats go over there.” Point to the right blue boat box. “Okay? So if it’s a rabbit, then it goes here.” Point to the left red rabbit box. “And if it’s a boat, then it
goes there.” Point to the right blue boat box. “This is the shape game. I’ll go first. Rabbits go here.” Place a blue rabbit card upside down in the left red rabbit box. “And boats go here.” Place a red boat card upside down in the right blue boat box.

**Pre-Switch Trials.** “Now it’s your turn. If it’s a rabbit, then put it here.” Point to the left red rabbit box. “But if it’s a boat, then put it there.” Point to the right blue boat box. “Here is a blue rabbit. Where does this go?” Wait for the child to place the card in one of the boxes. If the child is correct, say: “That’s right! Very good!” If the child is incorrect, say: “No, that’s not right. Remember the rules.” Proceed to the next trial.

“If it’s a rabbit, then put it here.” Point to the left red rabbit box. “But if it’s a boat, then put it there.” Point to the right blue boat box. “Here’s a red boat. Where does this go?” Wait for the child to provide an answer, and respond as described above. Proceed to the next trial.

“If it’s a rabbit, then put it here.” Point to the left red rabbit box. “But if it’s a boat, then put it there.” Point to the right blue boat box. “Here’s a blue boat. Where does this go?” Wait for the child to provide an answer, and respond as described above. Proceed to the next trial.

“If it’s a rabbit, then put it here.” Point to the left red rabbit box. “But if it’s a boat, then put it there.” Point to the right blue boat box. “Here’s a red rabbit. Where does this go?” Wait for the child to provide an answer, and respond as described above. Proceed to the next trial.
“If it’s a rabbit, then put it here.” Point to the left red rabbit box. “But if it’s a boat, then put it there.” Point to the right blue boat box. “Here’s a blue rabbit. Where does this go?” Wait for the child to provide an answer, and respond as described above. Proceed to the post-switch trials only after the child has responded correctly to 5 consecutive pre-switch trials.

Post-Switch Trials. “Now we are going to switch. We are not going to play the shape game anymore. We are going to play the colour game. When it is red, you have to put it in this box.” Point to the left red rabbit box. “But whenever it is blue, then it goes in that box.” Point to the right blue boat box. “We don't put red things in that box. No way. We put red things over here.” Point to the left red rabbit box. “And only when it's blue does it go over there.” Point to the right blue boat box. “If it's blue, then it goes there.” Point to the right blue boat box. “If it's red, then it goes here.” Point to the left red rabbit box.

“Okay, so remember - if it is red, then put it here,” Point to the left red rabbit box. “But if it is blue, put it there.” Point to the right blue boat box. “Here is a red boat. Where does this go?” Wait for the child’s response, and do not provide them with feedback. “Okay, let’s do another!”

“If it is red, then put it here,” Point to the left red rabbit box. “But if it is blue, put it there.” Point to the right blue boat box. “Here is a red rabbit. Where does this go?” Wait for the child’s response, and do not provide them with feedback. “Okay, let’s do another!”
“If it is red, then put it here,” Point to the left red rabbit box. “But if it is blue, put it there.” Point to the right blue boat box. “Here is a blue boat. Where does this go?” Wait for the child’s response, and do not provide them with feedback. “Okay, let’s do another!”

“If it is red, then put it here,” Point to the left red rabbit box. “But if it is blue, put it there.” Point to the right blue boat box. “Here is a red boat. Where does this go?” Wait for the child’s response, and do not provide them with feedback. “Okay, let’s do another!”

“If it is red, then put it here,” Point to the left red rabbit box. “But if it is blue, put it there.” Point to the right blue boat box. “Here is a blue rabbit. Where does this go?” Wait for the child’s response, and do not provide them with feedback.

“Okay, that was fun! Good job on that game!” The child receives a score of 1 for every post-switch trial that he/she answers correctly (maximum total = 5).
Appendix C
Children in the Home

1. Does your child [child’s name] live with other children?

☐ Yes ☐ No

If yes, please proceed with the questions below. If no, thank you for completing this questionnaire!

Please list the age and birthday of all siblings or other children living in [child’s name]’s home throughout his or her life.

<table>
<thead>
<tr>
<th>AGE:</th>
<th>BIRTHDAY:</th>
<th>SIBLING:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child 1:_______________________</td>
<td>_____________________</td>
<td>☐ Yes ☐ No</td>
</tr>
<tr>
<td>Child 2:_______________________</td>
<td>_____________________</td>
<td>☐ Yes ☐ No</td>
</tr>
<tr>
<td>Child 3:_______________________</td>
<td>_____________________</td>
<td>☐ Yes ☐ No</td>
</tr>
<tr>
<td>Child 4:_______________________</td>
<td>_____________________</td>
<td>☐ Yes ☐ No</td>
</tr>
<tr>
<td>Child 5:_______________________</td>
<td>_____________________</td>
<td>☐ Yes ☐ No</td>
</tr>
<tr>
<td>Child 6:_______________________</td>
<td>_____________________</td>
<td>☐ Yes ☐ No</td>
</tr>
</tbody>
</table>

Have the children listed above lived with [child’s name] for his/her entire life (or, in cases where the children listed are younger than [child’s name], since the children listed above were born?)

☐ Yes ☐ No

If no, please provide details on the time periods that each of the children listed above resided in the same home as [child’s name].
Appendix D

Theory of Mind Scale Task Scripts

Diverse Desires

The child is shown a stuffed animal (Big Bird) and a sheet of paper with an apple and a cookie depicted on it. “Here is Big Bird. It’s snack time, and Big Bird wants a snack to eat. Here are two different snacks: an apple and a cookie. Which snack would you like best? Would you like the apple or the cookie best?”

If the child chooses the apple, provide the following response: “Well, that’s a good choice, but Big Bird really likes cookies. He doesn’t like apples. What he likes best are cookies.” If the child chooses the cookie, provide the following response: “Well, that’s a good choice, but Big Bird really likes apples. He doesn’t like cookies. What he likes best are apples.”

“So, now it’s time to eat. Big Bird can only choose one snack, just one. Which snack will Big Bird choose? The apple or the cookie?” To pass the task, the child must respond to the target question with the food item that is in accordance with Big Bird’s desires.

Diverse Beliefs

The child is shown Cookie Monster and a sheet of paper with a garage and a tree depicted on it. “Here’s Cookie Monster. Cookie Monster wants to find his cat. His cat might be hiding in the tree, or it might be hiding in the garage. Where do you think that cat is? In the tree or the garage?”
If the child chooses the tree, provide the following response: “Well, that’s a good idea, but Cookie Monster thinks his cat is in the garage. He thinks his cat is in the garage.”

If the child chooses the garage, provide the following response: “Well, that is a good idea, but Cookie Monster thinks his cat is in the tree. He thinks his cat is in the tree.”

“So where will Cookie Monster look for his cat? In the tree or in the garage?” To pass the task, the child must respond to the target question with the location that is in accordance with Cookie Monster’s beliefs.

**Knowledge Access**

Show the child a jewelry box with a toy frog inside a closed drawer. “Here’s a drawer. What do you think is inside the drawer?” Open the drawer to show the child its contents. “Let’s see. Oh, it’s a frog!” Proceed to close the drawer. “Okay, so what’s in the drawer?” Wait for the child’s response. If he/she responds incorrectly, repeat the task from the beginning until he/she is able to provide the correct answer.

Place a stuffed animal (Kitty) on the table in front of the child. “This is Kitty. Kitty has never ever seen inside the drawer.” Test question: “So, does Kitty know what is in the drawer?” Memory question: “Did Kitty see inside the drawer?” To be correct, the child must answer “No” to the test question and “No” to the memory question.